UNIVERSITY OF VAASA

FACULTY OF TECHNOLOGY

TELECOMMUNICATIONS ENGINEERING

Phillip Babatunde Oni

ERGODIC CAPACITY AND EFFECTIVE CAPACITY OF SPECTRUM SHARING COGNITIVE RADIO WITH MRC OVER NAKAGAMI FADING: A COMPARATIVE ANALYSIS

Master's Thesis for the degree of Master of Science in Technology submitted for inspection in Vaasa, 29, November, 2012.

Supervisor   Professor Mohammed Salem Elmusrati

Instructor   M.Sc. (Tech.) Ruifeng Duan

ACKNOWLEDGEMENT

TABLE OF CONTENTS

ABBREVIATIONS

3G	Third Generations
4G	Fourth Generations
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
B_w	Bandwidth
CR	Cognitive Radio
CSI	Channel State Information
C_{erg}	Ergodic Capacity
DSANs	Dynamic Spectrum Access Networks
E_d	Magnitude of Signal Energy
E_C	Effective Capacity
FCC	Federal Communication Commission
IEEE	Institute of Electrical Electronics Engineers
ISM-RB	Industrial, Scientific, and Medical Radio Bands
LOS	Line of Sight
MRC	Maximal Ratio Combining
PDR	Programmable Digital Radio
P_i	Signal Instantaneous Power
PU_{rx}	Primary User Receiver
RF	Radio Frequency
RKRL	Radio Knowledge Representation Language
SDR	Software Defined Radio
SNR	Signal-to-Noise Ratio
SU_{tx}	Secondary User Transmitter

SU_{rx}	Secondary User Receiver
xG	NeXt Generation
WSN	Wireless Sensor Network
ISI	Intersymbol Interference
PDF	Probability Density Function
CDF	Cumulative Distribution Function
P_o	Probability of Outage
BPSK	Binary Phase Shift Keying
P_t	Transmit Power
Q	Interference Power Constraint

SYMBOLS

λ_T	Signal Energy Threshold
$\boldsymbol{P_d}$	Probability of Detection
$\boldsymbol{P_f}$	Probability of False Alarm
$\Gamma(.)$	Complete Gamma Function
$\Gamma(.,.)$	Incomplete Gamma Function
$Q_m()$	Generalized Marcum Q-function
f_o	Coherence Bandwidth
f_d	Doppler Frequency
f_c	Carrier Frequency
σ_τ	Root Mean Square Delay Spread
D_s	Doppler Spread
T_c	Coherence Time
T_s	Symbol Period or Symbol Time
B_c	Channel Bandwidth
B_s	Signal Bandwidth
$J_0(.)$	Bessel Function
$E\{.\}$	Expected Value Operator
$B(.,.)$	Beta Function
${}_2F_1(\dots)$	Gauss Hypergeometric Function

LIST OF FIGURES

LIST OF TABLES

UNIVERSITY OF VAASA
Faculty of Technology

Author:	Phillip Babatunde Oni
Topic of the Thesis:	Ergodic Capacity and Effective Capacity with MRC for Cognitive Radio over Nakagami Fading: A Comparative Analysis
Supervisor:	Professor Mohammed Salem Elmusrati
Instructor:	Ruifeng Duan
Degree:	Master of Science in Technology
Department:	Department of Computer Science
Degree Programme:	Master's Programme in Telecommunication Engineering
Major of Subject:	Telecommunications Engineering
Year of Entering the University:	2011
Year of Completing the Thesis:	2012 **Pages: 116**

ABSTRACT

The licensed spectrum is becoming more congested due to increase in number of mobile users and wireless applications. This increase in spectrum usage necessitates the need to efficiently use the underutilized spectrums. While spectrum allocated to wireless communication is becoming congested, other licensed and unlicensed spectrums are underutilized. In response to this underutilization, cognitive radio has been proposed to support efficient use of the spectrum. With cognitive radio, radio devices can dynamically sense and use idle spectrums (white spaces) using their autonomous detection capability based on different spectrum sharing techniques. These spectrum sharing techniques promote coexistence and cooperation among dissimilar wireless technologies. As with other radio technologies, signal propagation in cognitive radio experiences multipath effects and causes interference with other users. Hence, this thesis extensively investigated the system capacity of cognitive radios when the channel encounters Nakagami-m fading and the maximal ratio combining (MRC) antennas diversity is implemented at the secondary user.

The effective and ergodic capacity are mathematically and numerically analyzed and simulated. Therefore, this thesis covers the mathematical frameworks for analyzing the ergodic capacity and effective capacity of spectrum sharing cognitive radios with MRC antennas diversity under Nakagami fading. The maximum achievable information transmission rates at the physical layer (PHY) and the data link layer are obtained using the ergodic capacity and the effective capacity mathematical models respectively. The system capacity in each model scales as a logarithmic function of the channel power gains, subject to the average interference power and the delay quality of service (QoS) constraint in the case of effective capacity. The results obtained depict the maximum achievable information transmission rate when the secondary user (SU) is implemented with multiple antennas based on MRC diversity method, the channel fading statistically follows Nakagami-m distribution and the transmit power of the SU is subject to average interference power constraint to avoid harmful interference with the primary user (PU).

KEYWORDS: Cognitive radio, spectrum hole, fading, ergodic capacity, effective capacity, Rayleigh, Nakagami-m, white space, multipath, maximal ratio combining, underlay

1.   INTRODUCTION

Communication through air medium practically relies on the radio portion of the electromagnetic spectrum with the main objective of allowing users on the move, anywhere, anytime to communicate with other users easily and efficiently. This goal has been achieved, in fact, many research estimates proved and suggested that the adoption of wireless communication for different applications including mobile communication is increasing and will continue to expand as a result of the emergence of improved and new standards such as IEEE 802.15.4 (WSN), third generations (3G), fourth generation (4G), and a host of them expected to hit the market in the next decades, promising features for *triple-play* services (video, voice, and data), high throughput, capacity, and performance. Therefore, the concern is not whether these technologies and emerging standards will be useful, applicable, or adopted, but developing efficient access techniques that will foster peaceful coexistence of the these wireless technologies on the limited available radio frequency spectrum, especially through efficient spectrum allocation and usage.

Although wireless communication is growing with the advent of new standards, protocols, technologies, and its usage in different applications, but this beneficial growth introduces challenges, not only in the design of high capacity communication hardware, but also in the management and efficient use of the scarce radio frequency spectrum to accommodate this growth. Conventionally, concepts like *cellular design* have been proposed in the past to solve the problem of spectral congestion and user capacity constraints – with the goal of providing high capacity using limited spectral allocation available for mobile communication. Unfortunately, due to the exponential growth, these concepts are no longer sufficient to provide room for more users and applications (growth) in the mobile communication domain.

Therefore, the major problem has shifted from the design of economical hardware and systems to lack of frequency spectrum or bandwidth to handle the increasing demand. While mobile communication is confronted with spectrum availability constraint due to congestion and interference on its limited allocated frequency band, frequency spectra allocated to other services and applications are underutilized. In response to the need for efficient use of the radio spectrum, a spectrum access and sharing technique, *cognitive radio* (CR) has been proposed. Cognitive radio promises efficient use of the radio frequency spectra and peaceful coexistence of mobile

communication users with other primary users or dissimilar radio technologies on the intended shared spectrum – allowing users to transmit on any available frequency, including those not primarily allocated to them, anytime, anywhere while avoiding interference with other users, especially the primary or licensed users.

The fundamental concepts and architecture of cognitive radio will be thoroughly introduced in this text, starting with a brief discussion on frequency spectra allocation and evidence of spectrum underutilization. As an ideal preamble to the analysis of capacity for spectrum sharing CRs, a general discussion on the radio propagation phenomenon will be covered with focus on the statistical characteristics of three cases of flat fading, namely, *Nakagami-m, Rayleigh, and Ricean fading models* because it is important to have a clearer picture of the phenomena causing channel fading and signal degradation in wireless communication. Also, a chapter in this text has been dedicated to discussion on receiver diversity techniques in radio technologies, which fosters our understanding on the importance or benefits of implementing multiple antennas at the receiver input. Then subsequent chapter discusses the channel capacity models and presents the numerical results for analyzing the capacity of spectrum sharing cognitive radios.

The main idea of this thesis is centered on the analysis of the capacity of spectrum sharing cognitive radios with maximal ratio combining (MRC) diversity antennas at the secondary receiver under *Nakagami fading* along the path between the secondary transmitter (SU_{tx}) and the secondary receiver (SU_{rx}), and between the SU_{rx} and the primary receiver (PU_{rx}). In order to mathematically and numerically illustrate the maximum achievable information transmission rates of a spectrum sharing CR, two channel capacity models, namely, the ergodic capacity and effective capacity are considered and compared on the basis of the achievable information rate (*nats/s/Hz*) given an assumed number of MRC K antennas at the secondary receiver. It is also assumed that the Channel Side Information (CSI) is known at both the receiver and transmitter. This thesis will undoubtedly contribute to the ongoing research on spectrum sharing techniques and coexistence among dissimilar radio technologies. The next section discusses the issues surrounding frequency allocation and usage, and provides an insight into the efficient use of the radio spectrum through the implementation of spectrum sharing techniques suggested in cognitive radio.

2.   FREQUENCY SPECTRA ALLOCATION AND COGNITIVE RADIO

2.1.   FREQUENCY SPECTRA AND ALLOCATION

Radio communication and other wireless applications rely on certain licensed and unlicensed (ISM-RB) frequencies for propagation of signals from the transmitter to the receiver and vice versa. The radio portion of the electromagnetic spectrum is subject to license from an appointed government agency in charge of spectrum management such as Federal Communication Commission (FCC) in the United States and Office of Communication (OfCom) in UK. Licenses are issued based on spectrum availability. These bands of frequencies used for radio communication are technically referred to as frequency spectra. In order to prevent two or more wireless applications from interfering with one another during transmission, different frequencies are allocated to different radio applications, and this allocation is usually documented through a license that indicates the range of frequencies allowed per user (or application) for transmission. In most regions, frequency allocation depends on availability and planning; for example, some frequencies are reserved for scientific and medical research purposes while some are used for military applications and mobile communication.

Figure 1 shows an example of spectrum utilization chart. From this chart, it is apparent that certain wireless applications are growing in number and their allocated frequency spectra are getting congested. The dynamic increase in the access to the limited spectrum for mobile services in the recent years (Akyildiz et al., 2006: 2128) could be the reason for those heavily
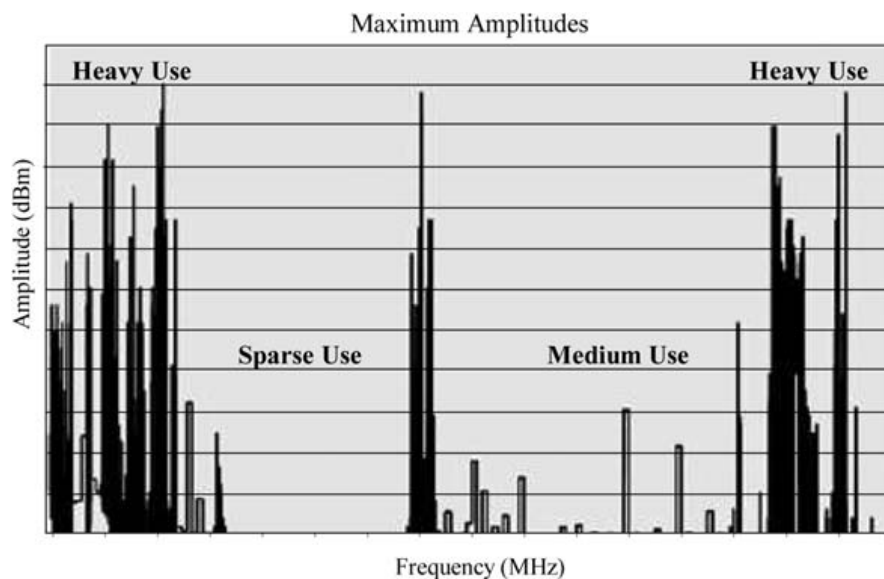
Figure 1. Spectrum Utilization (Akyildiz et al., 2006: 2128)

used spectra, which are probably those allocated to mobile communication. Similarly, the spectrum allocation chart published by the National Telecommunications and Information Administration provides clear overview of heavily used spectrum and the limited availability of spectrum[1] for future and emerging wireless applications. However, some radio applications other than mobile services do not make use of the allocated frequencies during certain time of the day and in certain geographical area, thereby causing the problem of spectrum underutilization, which need to be solved.

2.2.   COGNITIVE RADIO

Conventionally, license restricts the use of specific spectrums to an authorized primary user only, and no other user (*secondary*) is allowed to transmit or receive on the same frequencies in order to prevent interference with the legitimate (*primary*) user's transmission. This restriction on license-based spectrum causes underutilization of the electromagnetic spectrum in a situation whereby the band of frequencies allocated to primary users are not being used during a certain period of time and in specific geographical area – this phenomenon of unused frequencies has been described as *spectrum holes* (Haykin, 2005: 201). Considering the increase in number of wireless applications there is need for efficient use of the limited available spectrum in order to accommodate the growth. This inefficiency has prompted the development of an opportunistic based technique (Akyildiz et al., 2006: 2128) that allows radios to dynamically access the available spectrum.

Spectrum utilization could therefore be improved by allowing secondary users to access the spectrum at the time and location when such band of frequencies is not being used by the primary user (Haykin, 2005: 201). In response to this need, a renowned framework has been proposed to improve spectrum utilization by allowing the use of spectrum holes *when* and *where* they exist. This framework is known as cognitive radio (CR), which include software defined radio (SDR), whose initial focus was on radio knowledge representation language (RKRL) (Mitola, 2000: 1; Mitola et al., 1999: 13; Mitola, 1999: 32). As culled from Haykin (2005),

---

[1] Frequency Allocation Chart: National Telecommunications and Information Administration
http://www.ntia.doc.gov/files/ntia/publications/2003-allochrt.pdf

cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e. outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g. transmit-power, carrier-frequency, and modulation strategy) in real time, with the primary objectives of providing highly reliable communications whenever and wherever and efficient utilization of the radio spectrum (Haykin, 2005: 201-202). CR including SDR provides the possibility of implementing some level of computational intelligence in radio devices, so that such intelligent radios can dynamically detect unused spectrum and share the spectrum without interference with other primary or secondary users, support user communication needs by capturing/scanning for the best available spectrum, maintain seamless change to better available spectrum, and provide fair scheduling method. These dynamic or intelligent characteristics of CR devices are described in Akyildiz et al. (2006) as the main functions for cognitive radios in NeXt Generation (xG) Networks, which support *spectrum-aware* communication protocols (Akyildiz et al., 2006: 2128).

Spectrum Sensing:	Detecting unused spectrum and sharing the spectrum without harmful interference with other users.
Spectrum Management:	Capturing the best available spectrum to meet user communication requirements.
Spectrum Mobility:	Maintaining seamless communication requirements during transition to better spectrum.
Spectrum Sharing:	Providing the fair spectrum scheduling method among coexisting xG users.

Table 1. Required functionalities of CR (Akyildiz et al., 2006: 2128, 2131-2132)

Table 1 summarizes these functions expected from a CR enabled device. From Table 1 above, it can be inferred that radios adhering to CR computational capabilities will be able to determine which portion of the electromagnetic spectrum is available and detect the presence of the primary users (licensed users). When the available spectrum is detected, the best available channel is selected in a coordinated fashion with other users sharing the same spectrum. The most crucial

feature is the ability of the secondary user to vacate the channel when the presence of a primary user is detected. Consequently, the functionalities could be compressed into two main characteristics or capabilities, namely, *Cognitive capability* and *Reconfigurability* (Akyildiz et al., 2006: 2129; Haykin, 2005: 202). Basically, the cognitive capability defines the spectrum awareness feature of the radio, that is, the ability to capture and sense the information from its radio environment. The reconfigurability is the capability defined on the SDR platform (Haykin, 2005: 202), which enables CR device to be dynamically programmed (adapt) to the radio environment based on the sensed information – programming the CR to transmit and receive on a variety of frequencies and use different transmission access mechanisms based on the hardware configuration (Akyildiz et al., 2006: 2129). Reconfigurability basically deals with the radio's capability to dynamically adjust or tune its operating parameters (*operating frequency, modulation, transmission power, and communication technology*), (Akyildiz et al., 2006: 2129) for transmission anytime and anywhere, by considering the *spectrum-hole* information sensed from its operating RF environment, without any unforeseen modification to its hardware configuration.

2.2.1. COGNITIVE RADIO AS A FRAMEWORK

In order to enhance the effectiveness of service delivery, especially for xG wireless networks (Akyildiz et al., 2006: 2128), there is need to include certain computational capabilities that are currently missing in wireless nodes by embedding *model-based reasoning* framework (Mitola, 2000: 45). Cognitive radio is the proposed framework to extend this model-based reasoning concept, to include interaction with the radio RF environment. This computational interactive capability consist of three cognitive tasks (Haykin, 2005: 202) or steps (Akyildiz et al., 2006: 2131) as illustrated in **Figure** 2, which constitute a state diagram described as *cognitive cycle* (Mitola, 2000: 45; Akyildiz et al., 2006: 2132; Haykin, 2005: 202). The cognitive cycle in **Figure 2** states the computational tasks or steps that must be accomplished by a radio, for it to be cognitive in practice. The real time interaction with the RF environment (cognitive capability) is used by the radio to determine appropriate communication parameters for the purpose of adapting to the environment.
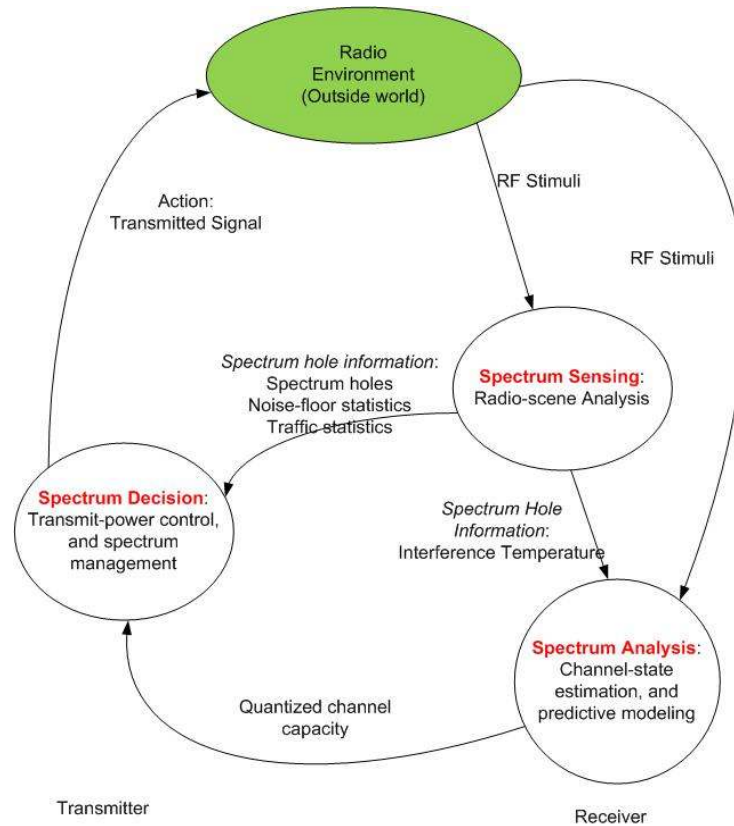
Figure 2. Cognitive Cycle (Haykin, 2005: 202; Akyildiz et al., 2006: 2131).

The cognitive capability (built on the three tasks defined in cognitive cycle) and the reconfigurability are two constituents of the cognitive radio framework. This framework is the proposition for building increasingly powerful computational intelligence capabilities (Mitola Joseph., 1999: 26) in xG radios. Such radios, in addition to their current flexibility nature, will need to implement cognitive ability, which will help their understanding and accessibility of their own radio structure – knowing basic facts about radio rather than just being radios without knowledge of their own structures (*awareness*); such as understanding equalizer's structure and functions. To achieve this computational model, the radio can use RKRL with its hardware design based on the proposed CR framework depicted in **Figure** 3, including the radio components as illustrated in the framework (Mitola Joseph., 1999: 26).
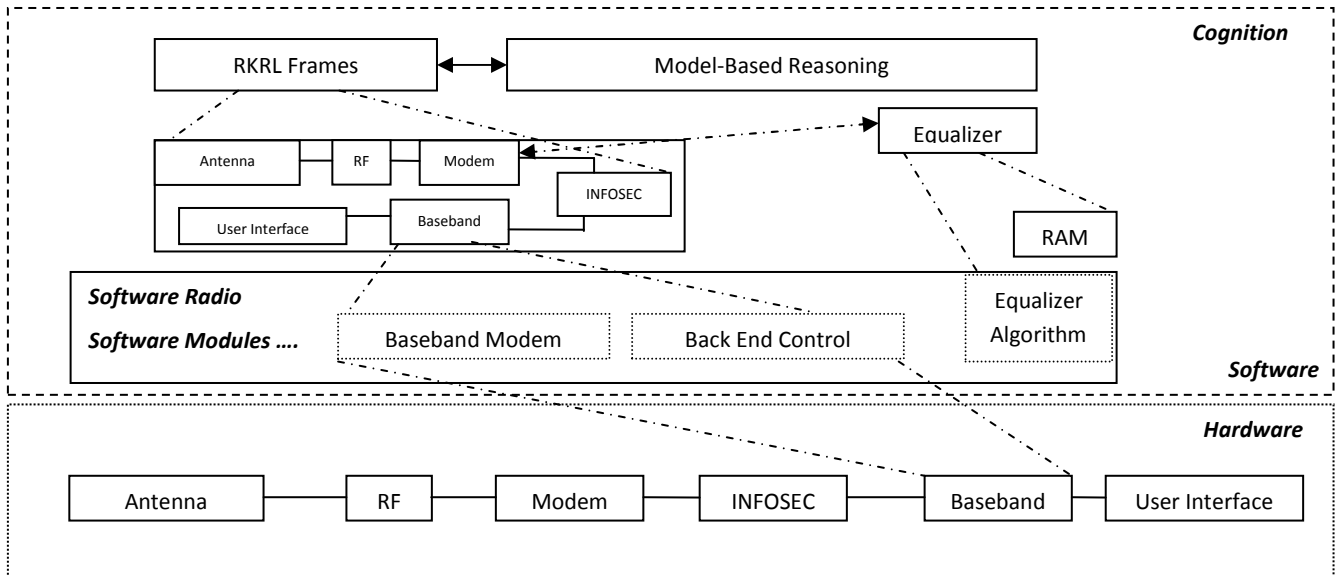
Figure 3. Cognitive Radio Framework (Mitola, 1999: 27; Mitola et al., 1999: 14)

The hardware unit of the framework (radio's model) contains a representation of radio's functions or components, namely antenna, RF conversion unit, INFOSEC modem, baseband processor, and user interface that will form a complete CR architecture (components of the radio transceiver) as shown in **Figure 4** and **5**, which is discussed later section *1.2.3*. In the software module, there is a baseband processor containing the software for the baseband modem and the equalizer module for protocol and control respectively. The equalizer algorithm provides a structural knowledge about the equalizers, providing information such as the taps' representation of the channel impulse (Mitola, 1999: 27; Mitola, 2000: 46). In digital communication, generally, equalizers with taps are used for channel equalization – filtering the received signal to correct for the non-constant channel gain or compensating the variance of the difference between the transmitted data and the signal at the output. In Mitola (2000), the framework also contains the model-based reasoning module, which, with the use of RKRL frames, provides radios with the control intelligence, that is, the cognitive ability (Mitola, 2000: 46-47).

In a concise form, the cognitive radio supports the use of temporary free or unused spectrum (*spectrum hole*). Hence, cognitive radio's ultimate goal is to sense, obtain, and use the best available spectrum through *cognitive capability* – reasoning and sensing and *reconfigurability* – adjustment of radio operating parameters to meet the requirements of the target or sensed

spectrum hole; in other words, sharing the licensed spectrum without interfering with the transmission of licensed users (primary users). The cognitive capability is a host of three stages as depicted in the cognitive cycle in **Figure** 2, which are *spectrum sensing, spectrum analysis, and spectrum decision*, while the reconfigurable parameters (Akyildiz et al., 2006: 2132) in *reconfigurability* include operating frequency, modulation scheme, transmitted power, and communication technology.

2.2.2. COGNITIVE RADIO CAPABILITY

As described in the previous section, it can be deduced that the cognitive capability and reconfigurability of radios will promote the use of the best available spectrum anywhere, anytime; which is the primary goal of CR/SDR implementation. Therefore, for a radio to be cognitive, it must be able to perform certain functional tasks and reconfigure its operating parameters for adaptation to the operating environment spectrum condition. The three main cognitive steps highlighted in the cognitive cycle shown **Figure 2**, are *spectrum sensing, spectrum analysis,* and *spectrum decision*. At s*pectrum sensing* stage, the CR monitors the available spectrum and captures all radio related information in order to detect the existence of *spectrum holes*. The characteristics of all detected *spectrum holes* are estimated during the *spectrum analysis* stage, and these estimations are used at the *spectrum decision* stage where the cognitive radio determines transmission mode (appropriate modulation schemes), data rate, and the bandwidth of the transmission. Following this decision, the appropriate spectrum band is selected based on the estimated spectrum characteristics and user requirements (Akyildiz et al., 2006: 2131 – 2132).

2.2.3. COGNITIVE RADIO TASKS

For cognitive radio capability or cognitive tasks, signal processing and machine-learning procedures are the fundamental building blocks of the implementation. Based on the cognitive cycle, each cognitive step described in the previous section, is made up one or more tasks. As shown in **Figure 2**, the spectrum sensing requires a computational task for radio scene analysis, spectrum decision includes transmit power control and spectrum management tasks, while the

tasks required for the spectrum analysis step are channel-state estimation and predictive modeling. The radio-scene analysis, channel estimation and predictive modeling are carried out in the receiver (Rx) while the transmit-power control and dynamic spectrum management are performed by the transmitter (Tx). Therefore, the cognitive module in the transmitter must synchronize with receiver's cognitive module, and this can be achieved with the help of a *feedback channel* between them (Haykin, 2005: 202). The following sections discuss the three main cognitive tasks in more details.

2.2.4. RECONFIGURABILITY CAPABILITY OF COGNITIVE RADIOS

The reconfigurability of cognitive radio will be based on software implementation, and that is what the SDR framework is set to achieve. Reconfigurability capability of cognitive radios deals with the ability of the radio to adjust its operating parameters to fit the hosting radio environment without any modification to its hardware components; all the reconfigurability functions are proposed to be implemented and performed in the software. The reconfigurable parameters that could be defined in the software include but not limited to operating frequency, modulation scheme, transmission power, and communication technology (Akyildiz et al., 2006: 2132). The main purpose of the reconfigurability functions is to allow cognitive radios, based on the current spectrum characteristics, switch to a different band through dynamic reconfiguration of the transmitter and receiver parameters, and selection of the appropriate communication protocol parameters and modulation schemes before and during transmission; these functions require efficient computation resources, which are defined in software radios.

Considering this need for reconfiguration of radio operating parameters, the reconfigurability capability is centered on SDR, which necessitates the implementation of *environment-aware* radio. In Mitola (2000), the concept of environment-aware computing for cognitive radio is that, computational entities in radios have awareness of their locations, users, networks, and the entire operating environment (Mitola, 2000:13). When radios are aware of their locations, users, networks and environment at large, they can tune their operating frequency and reconfigure other parameters based on the information gathered from the radio environment. Reconfiguring the operating frequency means that CRs can change their *frequency* based on the information about

the radio environment, and changing the *modulation scheme* to adapt to the user requirements and current channel conditions. *Transmission power* is another important parameter that requires computational reconfiguration. During transmission, CRs might be required to change the magnitude of their transmitted power from lower power to higher power or vice versa. CRs might be subject to transmission power reduction to avoid interference with other users sharing the same spectrum, most importantly, the *primary user* (Akyildiz et al., 2006: 2132).

2.2.5. Cognitive Radio Architecture

Cognitive radio could be a software radio or a programmable digital radio (PDR) (Mitola et al., 1999: 27). **Figure 4** and **5** show the proposed CR transceiver and RF/Analog front end architecture respectively. The two major components to achieve cognitive capability in CR are the RF front-end and the baseband processing unit.
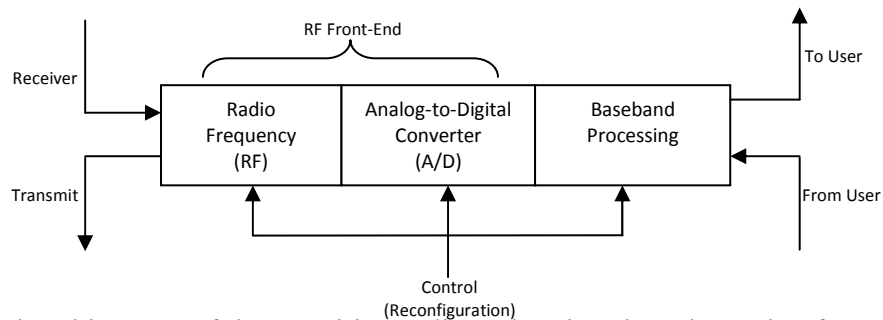
Figure 4. Physical architecture of the cognitive radio – showing the schematic of CR transceiver (Akyildiz et al., 2006: 2130)
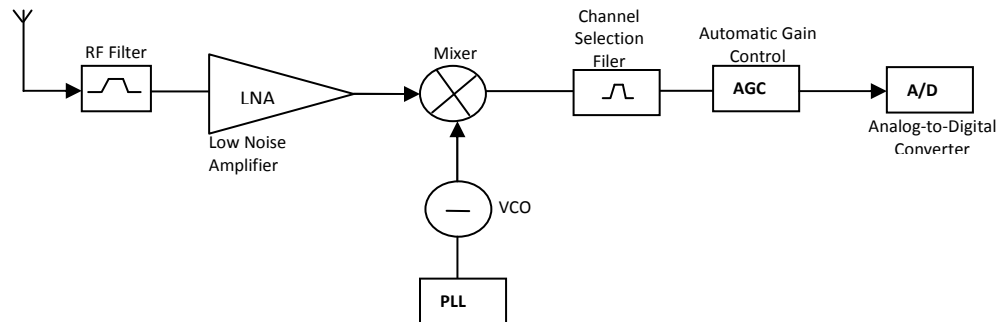
Figure 5. Physical architecture of the Cognitive radio showing block components of wideband RF/Analog front-end architecture (Akyildiz et al., 2006: 2130

The computation intelligence incorporated in CR will make reconfiguration of these components possible through a control bus, to adapt to the time-varying RF environment (Akyildiz et al.,

2006: 2130 – 2131). The CR components perform functions similar to that of existing radios, but the distinctive and novel feature of CR transceiver is the wideband sensing capability of the RF frontend – ability to sense and tune to a wide range of available frequency within its operating environment. According to Akyildiz et al. (2006), this wideband sensing function is mainly related to RF hardware technologies such as wideband antenna, power amplifier, and adaptive filter (Akyildiz et al., 2006: 2131). Having these components with wideband sensing capability implemented on RF hardware for CR will enable the radio to tune to any portion of a wide range of frequency spectrum within its operating environment, and provide real-time information on the operating frequency for cognitive decision.

The structure of CR wideband front-end shown in **Figure 5**, consists of hardware components with individual unique function. As culled from Akyildiz et al. (2006), the functions of each component are briefly discussed as follows:

1. *RF Filter*: generally, the objective of an RF filter is to sample an outgoing or incoming RF signal, block certain range of frequencies and allow the passage of other bands of frequency. Through the use of bandpass filtering, the RF filter can select the desired band and filtered out others. As Mitra (1999) explains filtering, filtering is used to pass certain frequency components in a signal through the system without any distortion and to block other frequency components (Mitra, 1999:5). This signal processing operation will be useful in CR operating environment where different band of frequencies exist. Bandpass filtering will passes only the frequency components between two cutoff frequencies. Therefore, the desired band from a wide range of frequency spectrum can be selected for operation.

2. *Low Noise Amplifier* (LNA): basically, LNA performs the amplification process by amplifying the for the purpose of minimizing the effect of noise. It is simply the multiplication of the signal by a constant (gain) greater than one. (Razavi, 2009)

3. *Voltage-controlled Oscillator* (VCO): this component generates a signal for a specified frequency range with a given voltage. This signal generated by the VCO is used in the mixer to convert the incoming bandpass signal to baseband or intermediate frequency band.

4. *Mixer*: when the received signal reaches the mixer, it is mixed with a local RF frequency generated by the VCO in order to convert the signal to the baseband or the intermediate frequency (IF). This mixing process produces a signal (IF) that is more convenient for detection and operation than the original received signal through channel selection stage.

5. *Phase Locked Loop* (PLL): the PLL is used to ensure that a signal remains on a specific frequency, preventing frequency deviation. In other words, PLL ensures that a signal is locked on a specific frequency.

6. *Automatic Gain Control* (AGC): this component prevents deviation in the power level or the gains of an amplified output signal; ensuring that the gain is constant over a wide range of input signals.

7. *Channel Selection Filter*: as the name implies, channel selection filter selects the desired frequency and rejects the adjacent channels. This channel selection filter could be a bandpass filter implemented in a superheterodyne radio receiver, and it could be a lowpass filer on a direct conversion receiver.

The design of CR RF-frontend architecture shown in **Figure 5** is currently challenging because the wideband signal received through the RF frontend needs to be sampled using high speed analog-to-digital (A/D) converter in order to perform measurement for detection of licensed user's (*primary*) transmission. Since the wideband RF antenna receives signals from wide range of transmitters operating at different power levels, bandwidths, and locations, the RF frontend must be implemented with the capability to detect low signals from a wide range of spectrum. The detection of these low signals requires A/D converter operating at multi-GHz speed, which is currently unrealizable (Akyildiz et al., 2006: 2131). Similarly, according to Razavi (2009), the design of broadband LNAs introduces tradeoffs between input matching, noise figure, gain, bandwidth, and voltage headroom (Razavi, 2009:391). The solutions suggested in Akyildiz et al. (2006) as alternative to the multi-GHz A/D requirement include reduction of the dynamic range of signal by filtering strong signals using tunable notch before A/D conversion, using multiple antennas to perform signal filtering in the spatial domain rather than in the frequency domain since multiple antennas can be used to receive signals selectively using different beamforming techniques (Akyildiz et al., 2006: 2131).

2.2.6. RADIO SCENE ANALYSIS AND SPECTRUM SENSING IN COGNITIVE RADIO

The detection of spectrum holes depends on the capability of CRs to analyze their radio operating environment. The spectrum sensing is an important task before spectrum sharing because it is the point in the cognitive cycle, **Figure 2**, where CRs can detect any available spectrum within their radio operating range that could be shared with the primary users – detecting the existence of spectrum holes or white spaces. According to Akyildiz (2006), the most efficient way to detect spectrum holes is to detect the primary users that are receiving data within the communication range of the CR users. This implies that before prospective secondary user can share a spectrum with the primary user, an analysis of the spectrum must be carried out to ensure that the spectrum is not currently being used by the primary user, and if the spectrum is unoccupied, the CRs enter the spectrum analysis and decision stages in **Figure 2** in order to determine the required operating parameters such as transmitting power, frequency, and modulation scheme.

The radio transceivers generate or emit RF stimuli while operating in their radio environment as illustrated in **Figure 2**, that are *nonstationary spatio-temporal* signals in nature because their statistical characteristics depend on time and space, and the radio scene analysis requires space-time processing, which comprises of two major functions, namely, spectral adaptive functions (estimation of interference temperature and detection of spectrum holes) and adaptive beamforming for interference control (Haykin, 2005: 204; Akyildiz et al., 2006: 2138). The main goal of spectrum sensing is centered on the fact that CR systems must be implemented in such a way that interference to primary users is avoided because the implementation and design of cognitive radios does not warrant the existing or legitimate spectrum owners (primary users) to change their radio architecture, infrastructure, or operating parameters, but the CRs that are secondary users must be sensitive to primary users' activity on the spectrum. There are different detection techniques that could be used for spectrum sensing, which are classified into three categories depicted in **Figure 6** as transmitter detection, cooperative detection, and interference-based detection. Similarly, these detection techniques are proposed as two compact algorithms classified as *Energy-based detection* and *feature-based detection* shown in Figure 7 (Wang, 2008 cited in Mekkanen, 2008: 65 – 66).
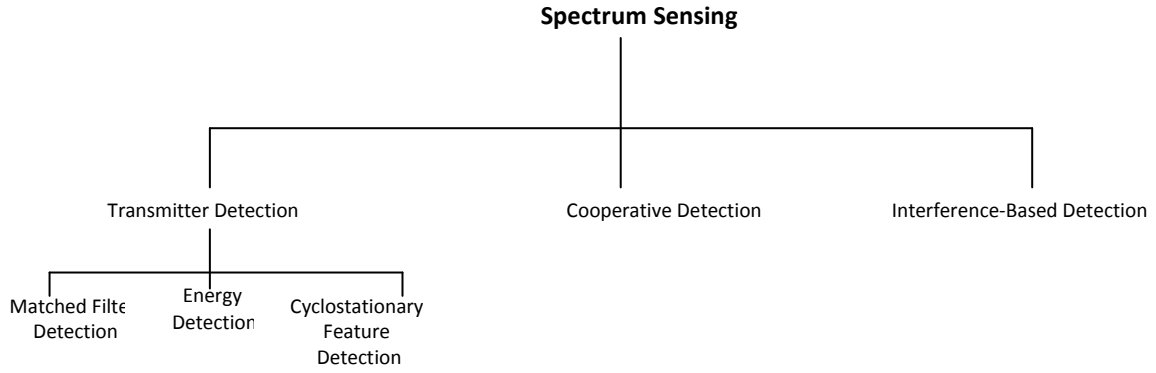
Spectrum Sensing

Figure 6. Classification of Spectrum Sensing Techniques (Akyildiz et al., 2006: 2138)
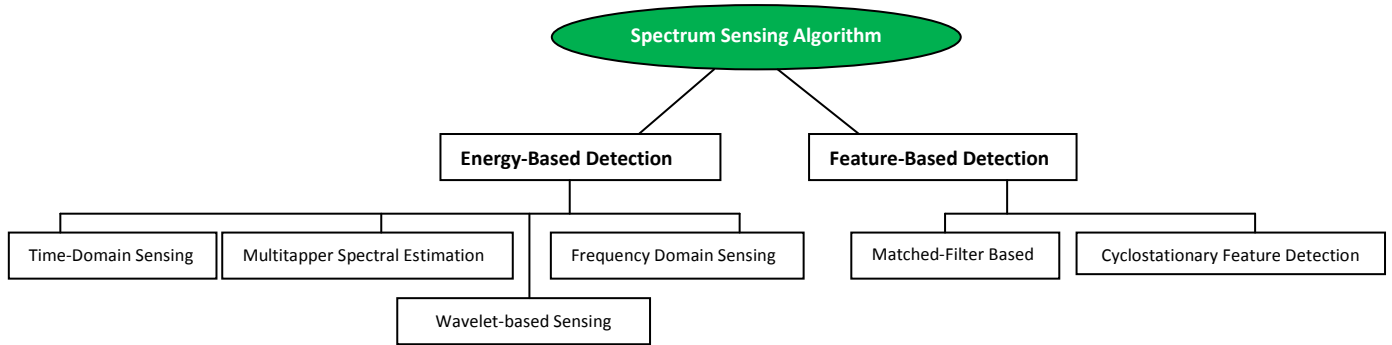
Figure 7. Algorithm-based Classification of Spectrum Sensing Techniques (Adapted from Wang, 2008 cited in Mekkanen, 2008: 72)

From the classification of spectrum sensing techniques in **Figure 6**, the **Transmitter detection** is a non-cooperative detection technique, which CRs could use to detect busy signal on any spectrum band purposely to determine if the spectrum is occupied or not. It is not sufficient to detect busy signal, so, CRs must also have the capability to determine if the detected signal is emitted by a primary transmitter in the spectrum. In this detection technique, secondary transmitter (cognitive radio) detects the weak signal from a primary transmitter by observing all spectrum users that locally existing on a certain spectrum. Under this technique, there three possible distinct methods, namely, *Matched Filter Detection, Energy Detection, and Cyclostationary Feature Detection* (Akyildiz et al., 2006: 2138). In general, the hypothesis model for defining transmitter detection technique is:

$$x(t) = \begin{cases} n(t) & H_0 \\ hs(t) + n(t) & H_1 \end{cases} \qquad (2.1)$$

From this model defined by equation (2.1), *x(t)* is the signal envelope received by the secondary user, *n(t)* is the additive white Gaussian noise (AWGN) while *h* is the channel amplitude gain. The three main methods under transmitter detection technique are centered on the hypothesis model given above. The H_0 is a null hypothesis, which indicates that there is no emitted signal from the licensed user (PU) or no PU is present in the sensed spectrum while H_1 is the opposite of hypothesis H_0, indicating the presence of a licensed user's transmission in the spectrum.

The **Matched Filter detection scheme** is used when the information about the PU's signal is known by the SU. Matched filter detection method uses coherence detection mechanism, and its main advantage is the possibility to achieve spectrum sensing with good performance index within a short sensing time through signal estimation using the received signal-to-noise ratio (SNR) as the parameter. Spectrum sensing with matched filter detection is less complex because most of the PU systems are implemented with pilot, spreading codes, or preamble, or training sequence for channel estimation and synchronization. This enable the SU (CRs) to estimate the received signal from the PU_{tx} but the disadvantage of this technique is that the estimation error becomes very high when the received SNR is low. In other words, if the estimation from the PU is not accurately known, the matched filter performance becomes low. Also, complexity arises when multiple PUs exist on the same radio scene with wide spectrum to be analyzed because the CRs need to estimate multiple primary signals. These complexity and constraints introduce research challenges in the design of matched filter for CR spectrum detection (Akyildiz et al., 2006: 2138, Liang et al., 2011: 3389).

Energy Detection Method is an alternative to matched filter detection, which can be used when the SU_{rx} is unable to gather adequate information about the PU signal; then the optimal detector is an energy detector when only the noise power is known to the SUs. Basically, in energy detection scheme, an energy detector or radiometer measures the received signal's energy level and compares it to a predefined energy level threshold. In order to determine the total energy in a received signal *x(t),* we defined *x(t)* as the integral of its instantaneous power $P_i(t)$ in the signal:

$$P_i(t) = |x(t)|^2 \tag{2.2}$$

$$E_d = \int_{-\infty}^{\infty} |x(t)|^2 \tag{2.3}$$

In other words, the output signal of the bandpass filter with bandwidth, B_w is squared and integrated over the sensing time (observation duration) interval T to obtain the magnitude of the received signal's energy, which is further compared with a threshold, λ_T to determine the presence of a PU in the spectrum. In the same context, assuming there are m sensed samples over time interval T, then the energy output E_d is given as thus:

$$E_d = \frac{2}{N_0} \sum_{i=n}^{m} |x(t)|^2 \tag{2.4}$$

Where E_d the energy output over m sensed samples (or number of observations), $x(t)$ is the received complex value discrete time signal, and N_o is the noise power. In order to make the energy detection decision, E_d is compared with the threshold, λ_T using:

$$x(t) = \begin{cases} H_1 \ if \ E_d > \lambda \\ H_0 \ \ if \ E_d < \lambda \end{cases} \tag{2.5}$$

Similar to the hypotheses discussed under (2.1), if magnitude of energy detected, is greater than the predefined threshold, it means hypothesis H_1 is true and a licensed user (PU) is present in the spectrum, but if the magnitude of energy detected E_d is less than the threshold value, λ_T it means that hypothesis H_0 becomes true and it indicates that no PU is detected in the spectrum. (Akyildiz et al., 2006: 2138 – 2139, Liang et al., 2011: 3389, Lundén, 2009: 17). The application of energy detection scheme in a *non-fading* radio scene (environment) where there is channel amplitude gain, h showing in (1.1), the probability of detection P_d and the probability of false alarm P_f could be modeled as follows:

$$P_d = P\{Y > \lambda \mid H_1\} = Q_m\left(\sqrt{2\gamma}, \sqrt{\lambda}\right) \tag{2.6}$$

$$P_f = P\{Y > \lambda \mid H_0\} = \frac{\Gamma\left(m, \lambda/2\right)}{\Gamma(m)} \tag{2.7}$$

Where γ is the SNR, $\Gamma(.)$ and $\Gamma(.,.)$ are complete and incomplete gamma functions Incomplete while $Q_m()$ is the generalized Marcum Q-function.

Under this non-fading condition or function simplified in equations (2.6) and (2.7), a low P_d means there is high probability that the CRs will not be able to detect the presence of a PU, and

this will increase the possibility of interference to the PU from SUs. In other words, the probability of detection in designing the detection threshold λ should be defined high enough to protect the PUs. Similarly, higher P_f will result in low spectrum utilization since false alarm increases the number of missed opportunities (untapped spectrum holes). Similarly as expressed in Liang et al., (2011), if both the signal and noise are real values that follow Gaussian distributions with zero mean and independent from one sample to another, the probability of detection and the probability of false alarm could be modeled respectively as:

$$P_d = P\{Y > \lambda \mid H_1\} = \frac{\Gamma\left(m/2, \lambda/(2+2\gamma)\right)}{\Gamma(m/2)} \tag{2.8}$$

$$P_f = P\{Y > \lambda \mid H_0\} = \frac{\Gamma\left(m/2, \lambda/2\right)}{\Gamma(m/2)} \tag{2.9}$$

Then the two hypotheses (or test cases) in equation (2.5) based on the Gaussian distributions with zero mean renders the energy detector test statistics to obey the following distribution function:

$$H_0 : Y_m \sim X^2_{2m} \tag{2.10}$$

$$H_1 : Y_m \sim X^2_{2m}(2\gamma) \tag{2.11}$$

From H_0 hypothesis, the detection test statistics follows central chi-square distribution with *2m* degrees of freedom while for H_1, it follows a non-central chi-square distribution with *2m* degrees of freedom and non-centrality parameter *2γ*. Contrary to the non-fading condition, the shadowing and the multipath fading factors should ideally be considered in energy detector design since it is most likely that most radios operate, in worst case scenario, under fading effect. Considering these factors, if P_f is independent of Γ and channel amplitude gain h fluctuates due to the shadowing and fading effects, then the probability of detection, P_d, depends on the instantaneous SNR with the probability function in (2.12) where f(x) is the probability distribution function of SNR under fading (Lundén, 2009: 17; Liang et al., 2011: 3389; Akyildiz et al., 2006: 2139).

$$P_d = \int_x Q_m\left(\sqrt{2\gamma}, \sqrt{\lambda}\right) f_\gamma(x) \ dx \tag{2.12}$$

In principle, **Figure 8a** and **8b** depict the block diagram of energy detector implementation. Although energy detection scheme is easy to implement and previous implementations of PUs detection in has been based on this technique, but there are drawbacks with possible solutions that might introduce some overheads in the practical design of the detector.
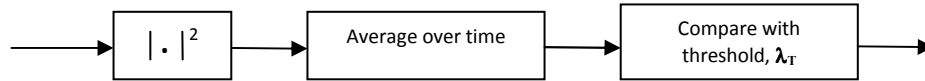
Figure 8a. Energy detector (radiometer) – Time domain
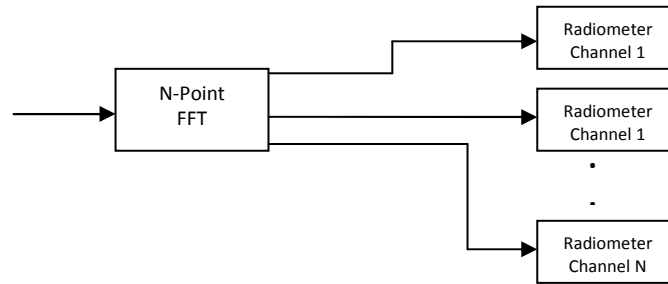
Figure 8b. Energy detector in Frequency domain (Lundén, 2009: 18)

The first drawback of energy detector is the performance constraint – as a result of uncertainty in noise power, obtaining accurate energy detection value might be difficult. Implementation of *pilot tone* in PU_{tx} has been proposed as the solution for the noise power uncertainty. An example of such implementation is the pilot energy detection for ATSC signals where the received signal is band-pass filtered around the pilot frequency then the energy detection is performed in frequency domain to obtain the hypothesis statistic as a function of the maximum squared fast Fourier transform (FFT), which is the compared to the predefined threshold, λ.

One more problem with energy detector is the ability to detect only the presence of a signal without an ability to differentiate signals types, which means there is possibility of detecting unwanted signals, leading to high probability of false alarm when the received signal contains unintended signals and finding spectrum becomes difficult. A novel solution to this problem is the *cyclostationary feature detection technique*, which has the capability to differentiate signals types (Lundén, 2009: 19; Liang et al., 2011: 3389; Akyildiz et al., 2006: 2139).

In **Cyclostationary feature detection**, the intention is to exploit the cyclostationary feature of modulated signals since modulated signals usually have statistical properties that vary cyclically with time (i.e. having multiple interleaving stationary processes). It is an alternative detection method to energy detection and matched filter detection schemes. In principle, this detection scheme can be obtained mathematically by analyzing the spectrum correlation function as thus:

$$R_y^\alpha(\tau) = E\left[y(n)y * (n - \tau)e^{-j2\pi\alpha n}\right] \tag{2.13}$$

While the cyclic spectral density (CSD) function of the received signal is as follows and α represents the cyclic frequency parameter:

$$S(f,\alpha) = \sum_{\tau=-\infty}^{\infty} R_y^\alpha(\tau)e^{-j2\pi f\tau} \tag{2.14}$$

This cyclostationary feature of modulated signal results from the fact that generally, modulated signals are composites of sine wave carriers, hopping sequence, spreading codes, pulse train, or cyclic prefixes, which contribute to the periodicity of such signals and having their mean and autocorrelation exhibit periodicity.

Although, the implementation of spectrum sensing with cyclostationary method is computationally complex, but it provides accurate detection estimation because of its ability to differentiate primary signal from the noise and other interference signals using their respective cyclic frequency as the basis for signal discrimination. Similarly, unlike energy detection scheme that is conditioned on instantaneous SNR under fading, cyclostationary scheme performance does not vary with SNR fluctuation. Therefore, its detection capability is reliable even in low-SNR radio environment. **Figure 9** depicts the basic implementation of cyclostationary sensing method (Akyildiz et al., 2006: 2139, Mekkanen, 2008:82; Liang et al., 2011: 3389 – 3390).
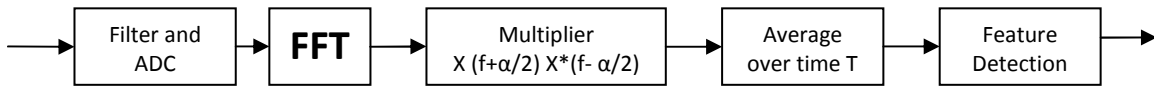
Figure 9. Implementation of Cyclostationary detection method (Mekkanen, 2008:82).

In general, the principle of sensing could either be *indirect sensing boundary* (transmitter detection) or *direct sensing boundary* (receiver detection). The spectrum detection techniques discussed so far fall under the transmitter detection model (also known as non-cooperative) as

depicted in **Figure 6**. There are downsides of transmitted detection model, which are the issue of *hidden node* and the inability of CR users to avoid the interference due to lack of knowledge about the PU_{rx} transmission and operating information as shown in **Figure 10**. Therefore, the lack of the sensing information from other users hinders the accuracy of spectrum detection (Akyildiz et al., 2006: 2139; Liang et al., 2011: 3389).
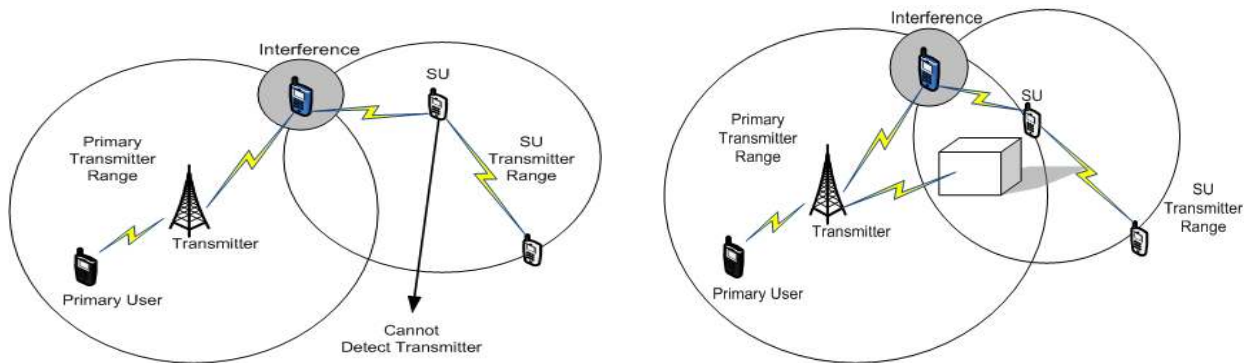
Figure 10. Transmitter Detection Problem: (a) Receiver uncertainty (b) Shadowing uncertainty
(Akyildiz et al., 2006: 2140)

Contrary to indirect spectrum sensing, direct spectrum detection provides a mechanism to protect the PU_{tx} from unwanted interference, which can be achieved by directly detecting the Rx (PU_{rx}) rather than the transmitter as in indirect spectrum detection. **Cooperative detection technique** is an example of direct sensing method, which is intended to solve the hidden node and interference problems. In cooperative spectrum sensing, the radio environment information from multiple SUs is incorporated together for the purpose of achieving accurate and improved sensing performance because it minimizes the uncertainty in a single user's detection process. As with other radio technologies, multipath fading and shadowing effects are severe sources of impairments and performance degradation in spectrum detection techniques – hindering accurate detection of the presence or absence of Pus (Akyildiz et al., 2006: 2140).

In order to primarily mitigate the effects of fading and shadowing in spectrum sensing techniques, cooperative detection methods are used to improve the detection probability in a severe shadowed environment. There are two forms of implementing cooperative sensing. It

could be implemented as a centralized sensing system where the base station gathers all the sensing information from all the SUs using some signal combining techniques or algorithms in order to accurate detect spectrum holes. Another form of implementation is distributed sensing system, which requires consistent exchange of spectrum information between cooperating SUs. Although, cooperative detection methods prevent the multipath fading and shadowing effects and provide accurate sensing information, but the main problem is the addition of certain operating parameters and coordinating algorithms which lead to transmission of overhead traffic on the resource-constrained networks. Also, the primary receiver's location uncertainty problem remains unsolved (Akyildiz et al., 2006: 2140).

Interference-based detection is another spectrum detection mechanism, which depends on the magnitude of measured interference. In an ideal radio operating environment, the transmitter is responsible for interference control through radiated power control, distance between individual transmitter (or location), and the out-of-band emissions. However, transmitters are not the only sources of interference; receivers can also contribute to RF noise floor, that is, interference can also occur at the receivers in a radio operating environment as shown in **Figure 10**. This fact renders the measurement of interference from the perspective of transmitters (transmitter-centric approach) insufficient due to accuracy issue.
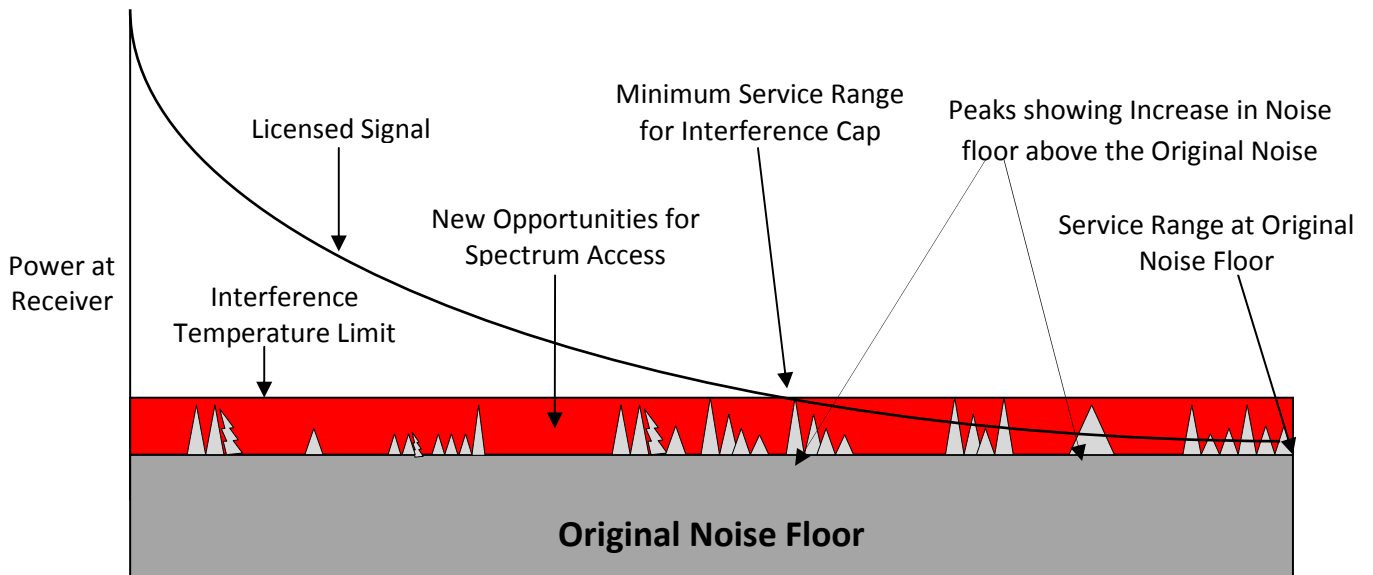
Figure 11. FCC Proposed Interference Temperature Model (Akyildiz et al., 2006: 2141)

The new metric for measuring noise floor has been proposed as *interference temperature* with a reference model developed by the FCC as illustrated in **Figure 11**. This approach suggests the measurement of interference as a real-time operation between the transmitter and the receiver in an adaptive manner as oppose to the transmitter-centric approach that supports the assessment of interference based on fixed operation in the transmitter (Akyildiz et al., 2006: 2140; Haykin, 2005; 203). The transmitter-centric technique is based on the design of transmitted power to approach a noise floor threshold at a certain distance from the transmitter. This does not provide assessment accuracy as a result of its non-adaptive nature to changes, such as signal degradation, in radio operating environment. On the other hand, interference temperature approach is an adaptive scheme that provides a "worst-case" measurement by managing interference at the receiver through estimation of the cumulative RF power from multiple transmissions. From the model illustrated in **Figure 11**, a radio station transmits signal in a range so that the received power approaches the original noise floor while at some points, the noise floor increases as more interfering transmissions are detected in the radio operating scene, as shown with peaks in the red region representing the interference temperature limit (Akyildiz et al., 2006: 2140; Haykin, 2005; 203).

This receiver-centric model is recommended for two reasons. Firstly, it provides accurate measurement of acceptable level of RF interference in a specific frequency band. Therefore, any unwanted transmission in this band is considered as interference and harmful to serviced users' transmissions because it increases the noise floor above the interference temperature limit. Secondly, if the measured aggregate interference from other transmissions (secondary users) does not exceed the interference-temperature limit, the frequency band could be made available to secondary users enforcing the interference temperature limit as the basis ("cap") for maximum power transmission allowed within the band (Haykin, 2005; 205). Hence, since cognitive radio is primarily a receiver-centric radio technology, the receiver requires an accurate and reliable spectral estimation of the interference temperature for the purpose of detecting spectrum holes.

The interference in the context of cognitive radio is the magnitude of service disruption caused by the secondary users to the primary users; in other words, as defined in Akyildiz et al. (2006) it is the expected fraction of primary users with service disrupted by the xG (cognitive radio)

operations. Therefore, for as long as the SUs transmissions do not result in interference above the interference-temperature cap, the spectrum can still be used by the SUs. However, this model introduces certain problems. It considers only the interference caused by a single SU, and does not consider radio environment with multiple SUs. Also, the SUs need to be aware of PUs' location before interference-temperature estimation could be performed, hence, when the location of PUs is unknown to the SU, estimation cannot be taken using this method. More so, other factors such as type of unlicensed signal modulation, antennas, ability to detect active licensed channels, power control, and activity levels of the licensed and unlicensed users must be considered, which result to design complexities when using interference-temperature method for spectrum detection (Akyildiz et al., 2006: 2140 – 2141).

Interference temperature estimation, as suggested in Haykin (2005), could be achieved in two ways. First, by estimating the power spectrum of the interference temperature resulted from the cumulative distribution of both internal sources of noise and external sources of RF energy using *multitaper method*, which produces near-optimal estimation. The second method as suggested, is the use of large number of sensors deployed in an RF environment to sniff RF transmissions, but depending on the RF environment (indoor or outdoor), large number of sensors might be needed for estimation that takes spatial variation of RF stimuli into consideration. The use of multiple sensors is possible in cognitive radio as a direct receiver detection method by exploiting the LO leakage power emitted by the RF front-end (Figure 5) of the PU_{rx}. These sensor nodes are mounted close to the primary receivers to detect the leakage LO power in order to determine the channel being used by the PUrx, and this information can be used by SUs (unlicensed users) to determine the possible spectrum holes (Haykin, 2005:205; Akyildiz et al., 2006: 2141).

2.2.7   Spectrum Sharing, Coexistence, and Cognitive Radio Paradigms

Following the spectrum sensing methodology discussed earlier, the detected spectrum holes will become available for CRs to share with the primary users (licensed users) using some certain spectrum sharing techniques discussed in this section – that is, the exploitation of spectrum holes. For the realization of dynamic spectrum access, interactions between different network components need to be implemented through cross-layer design approach. In cognitive radio, the most crucial interaction is that between *spectrum sensing* and *spectrum sharing* processes to

ensure spectrum efficiency, this interaction is shown in **Figure 12** where sensing information obtained from spectrum sensing at the Physical layer (PHY) is used for efficient spectrum sharing at the Link layer. Therefore, cross-layer design approach will play an important role in cognitive radio network design, to define algorithm for cooperation between spectrum sensing and spectrum sharing.
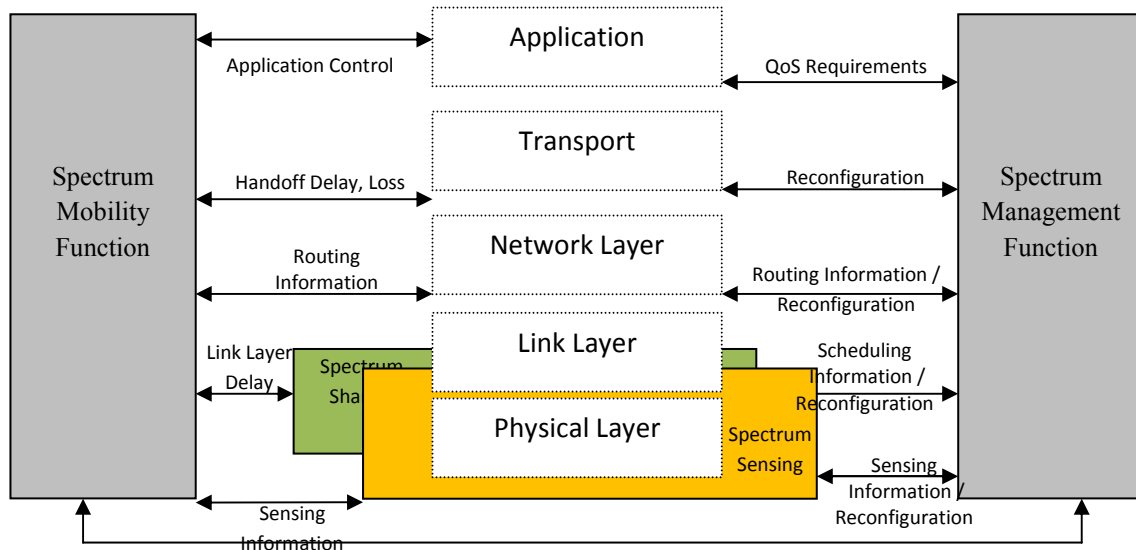
Figure 12. Handoff Decision, Current and Candidate Spectrum Information (Akyildiz et al., 2006: 2129).

Similarly, on the cross layer architecture for cognitive radio, **Figure 13** shows the required PHY and medium access control (MAC) functions for CR spectrum sensing that facilitate fair and efficient spectrum sharing.
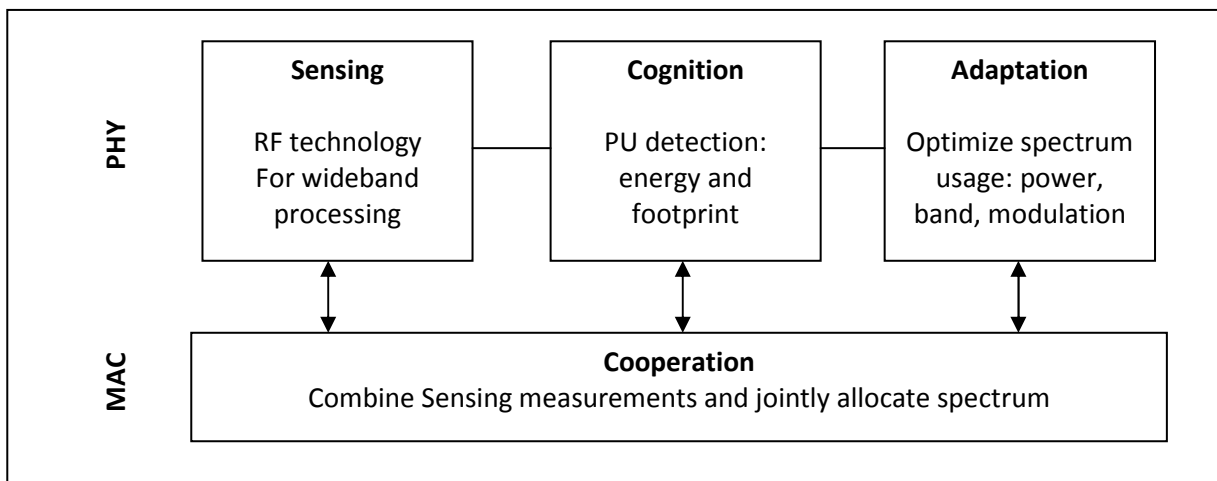
Figure 13. Cross Layer functionalities for Spectrum sensing cognitive radio (Cabric D. et al. 2004: 773)

The spectrum sharing mechanism is aimed at ensuring fair spectrum scheduling method among coexisting secondary users. Cognitive radio spectrum sharing is similar to medium access control in existing communication networks, and it remains one of the main challenges in open spectrum usage proposed in cognitive radio (Akyildiz et al., 2006; 2145). **Figure 14** depicts a simple radio operating environment with detected spectrum holes, which are available for exploitation by secondary or unlicensed users but there is need to define an efficient spectrum sharing method.
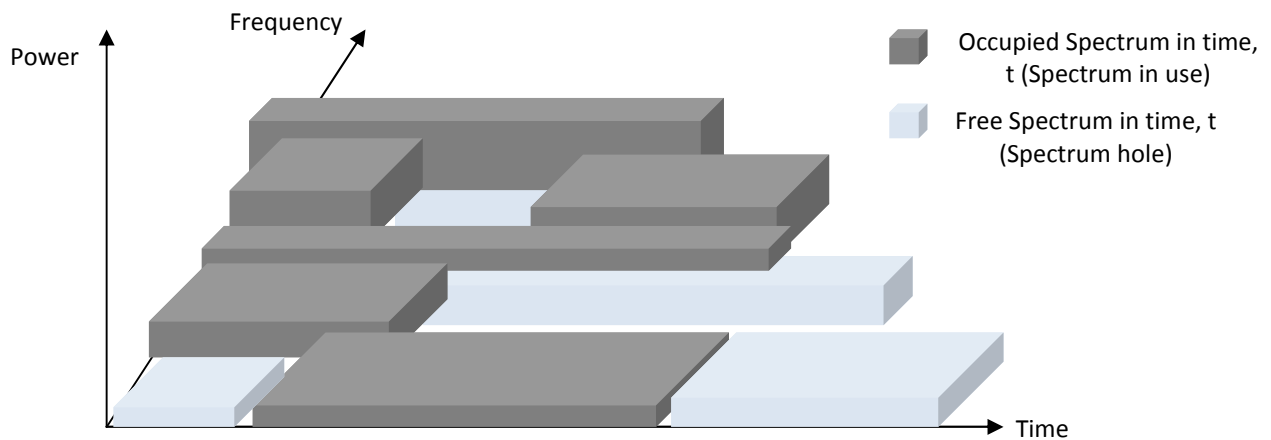
Figure 14. The Concept of Spectrum hole (Akyildiz et al., 2006: 2130)

There are five processes that constitute CR spectrum sharing, which are briefly illustrated in **Figure 15** as a top-down model from the initial step to the last. The spectrum sensing method has been discussed in previous section, allows CR users to be aware of their radio operating environment for the purpose of detecting available spectrum holes. The spectrum allocation step deals with the allocation of spectrum to CR users based on availability and spectrum usage policies. In order to avoid collision among multiple radios in an overlapping portion of the allocated spectrum, spectrum access mechanism must be defined as another step in spectrum sharing CR. After spectrum has been allocated to SU, there is need for the SU to inform the receiver about the selected spectrum to be used for transmission; that is what mandated the implementation of Tx-Rx handshake protocol. The final process required in spectrum sharing CR is the spectrum mobility needed at any point in time, to migrate the CR users to another vacant spectrum, when the current spectrum is being demanded by the primary user (licensed user)

Spectrum sharing techniques could be classified based on the architecture, spectrum access behavior, and access technology (paradigms) (Akyildiz et al., 2006: 2145 – 2146; Goldsmith et al., 2009:896).
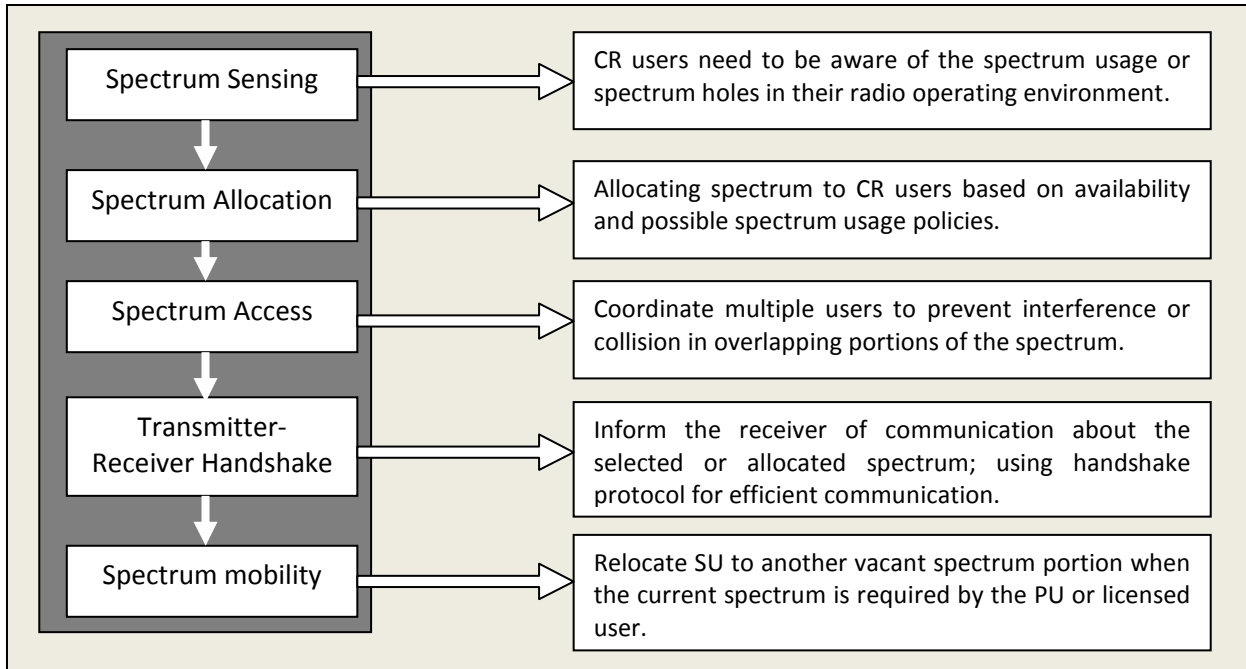
Figure 15. Spectrum sharing CR steps

Spectrum sharing in cognitive radio is similar to the medium access control (MAC) in other existing wireless/radio technologies but with different protocol design complexity and challenges. The challenges and complexity of CR spectrum sharing are unique to the MAC functions of other existing radios because of the issue of coexistence with licensed users and the wide range of available spectrum. With the introduction of CR, spectrum allocation is no longer restricted to licensed and unlicensed models because it promotes cognitive communication with noncognitive radios and dynamic spectrum usage with minimal interference or disruption with existing primary users. This advance radio technology requires efficient spectrum sharing techniques depending on the type of available network information or access technique, regulatory policies, architecture, and allocation behavior (Akyildiz et al., 2006: 2145; Goldsmith et al., 2009:895).

Depending on the network architecture, cognitive radios can share spectrum with primary users using centralized and distributed spectrum sharing techniques. As with other traditional centralized control/management system, in **centralized spectrum sharing**, where implementation of an infrastructure based network is feasible, a centralized entity coordinates the spectrum allocation and sharing processes. The main issue with this type of sharing technique is the service downtime when the central node experience failure causing the inability of the CRs to access the spectrum holes. In **distributed spectrum sharing**, each cognitive radio on the network is responsible for spectrum sensing information and measurement such as interference-temperature limit necessary for spectrum allocation and access. This technique is useful in a radio network environment where the deployment of infrastructure based network is not feasible, so each CR must have sensor attached to it as previously discussed (Akyildiz et al., 2006: 2146). More so, spectrum sharing techniques can be classified based on certain access behavior. The spectrum sharing techniques based on cognitive radios behavior could be cooperative or non-cooperative spectrum sharing technique. The **cooperative spectrum sharing** is closely related to centralize spectrum sharing, where each CR user on the network shares its spectrum sensing information such as the interference temperature measurement with other users in a cooperative manner, and this information is used by the spectrum allocation algorithm to make allocation decision. On the contrary, in **non-cooperative spectrum sharing** or *selfish technique*, each node does not share spectrum sensing information with one another (Akyildiz et al., 2006: 2146 – 2147; Peh and Ying-Chang, 2007:27).

These two spectrum sharing techniques could be compared on the basis of their spectrum utilization, fairness, and throughput capability. In Akyildiz et al., (2006), it is stated that non-cooperative solutions or algorithms may result in reduced spectrum utilization as a result of minimal communication requirements among nodes. This implies that non-cooperative technique is more bandwidth friendly, but there are more technical issues in CR that prevent its usage. As previously discussed, in spectrum sensing, there are two important probabilities, which are the probability of detection, P_d and the probability of false of alarm, P_f. The probability of detection indicates the probability of detecting PUs when the primary users are active, while the probability of false alarm deals with the probability of detecting the primary users when they are not active. The use of spectrum holes (unoccupied spectrum) by SU When the probability of

false alarm is high is not feasible because the SU will still have to vacate the channel when there is no PU (Akyildiz et al., 2006: 2147; Liang et al., 2011: 3389; Peh & Ying-Chang, 2007:27).

In (Peh & Ying-Chang, 2007), the performance of **cooperative spectrum sharing** has been investigated using these two probabilities. Cooperative technique simply collects all sensing data from all distributed SUs and makes spectrum allocation decision based on the analysis of such data. In their analysis, it was concluded that cooperating all SUs in the network does not achieve the optimum P_f or P_d, but coordinating group of users with higher PUs SNR, γ will yield the optimum probabilities; proved that cooperative technique can improve the probabilities of detection and false alarm (Peh and Ying-Chang; 2007; 27 – 32).

Most importantly, spectrum sharing techniques in cognitive radio, based on the access technology or the type of available network information and regulatory policies, can be categorized as *underlay, overlay*, and *interweave*. In other words, cognitive radios can underlay, overlay, or interweave their transmissions with existing radios or primary users with minimum or no interference. These spectrum sharing schemes are discussed in Goldsmith et al. (2009) as cognitive radio network paradigms. In underlay paradigm, CR users can coexist with primary users (licensed users) if the magnitude of interference caused to non-cognitive radios is below the predefined threshold. Overlay spectrum sharing scheme allows cognitive radios to use sophisticated signal processing and coding techniques to improve the communication of primary users, and obtain the free bandwidth for their own communication. Interweave spectrum sharing scheme is an opportunistic access technique, whereby cognitive radios opportunistically exploit spectrum holes for their transmissions without causing interference to other cognitive and non-cognitive radio transmissions. The following subsections throw more light to these paradigms (Akyildiz et al., 2006: 2147; Goldsmith et al., 2009:896).

1.   *Underlay Spectrum Sharing*: in this spectrum sharing or coexistence paradigm, the algorithm implements a technique that allows cognitive radios to be aware of the interference caused by their transmitters to the receivers of all noncognitive radios. Such awareness or cognitive capability can be implemented in CRs using the interference temperature model illustrated in **Figure 11**. This spectrum sharing technique is similar to

the spectrum access technique used in existing cellular network, whereby the transmit power of CRs become noise to the licensed users at certain portion of the spectrum allocation map of the radio operating environment. Underlay coexistence relies on certain transmit power control algorithms and sophisticated spread spectrum techniques to curtail interference or improve PU receiver's resistance to interference. In a simple term, cognitive radios and licensed users can concurrently transmit signals in the same radio operating environment if the CRs can keep the interference caused by their transmitters below certain threshold (Akyildiz et al., 2006: 2147; Goldsmith et al., 2009:896).

2.   *Overlay Spectrum Sharing*: this coexistence technique involves the use of certain signal processing and coding techniques by the CRs to improve the transmission of primary users. Improving the transmission of PUs means improved and efficient bandwidth utilization, and CRs can have the opportunity to obtain additional bandwidth for their own transmissions. In order for CRs to contribute to the noncognitive users' transmission improvement, they must have the knowledge of PUs' codebooks and messages. One way to obtain such knowledge is to be aware of the uniform coding standard being used by the PUs, and another way is to obtain it from the periodic broadcast of the codebooks by the PUs. Cognitive radios can therefore utilize this information and assign part of their power to transmit their own signals while assisting the noncognitive radios with the remaining part of their power for effective communication and peaceful coexistence (Goldsmith et al., 2009:896).

3.   *Interweave Spectrum Sharing*: This CR spectrum sharing technique is the fundamental motivation for cognitive radio networks because it is centered on the phenomenon of opportunistic communication. This spectrum sharing idea was postulated following the report submitted by the Spectrum Efficiency Working Group (SEWG) of FCC, concluding that a major part of the spectrum in both licensed and unlicensed bands is not utilized most of the time, even at some geographical location, so, there exist space-time-frequency (or spectrum holes) as also shown previously in Figure 14, which can be exploited by cognitive radios for their transmissions. It is opportunistic because cognitive radios are required to sense the activity or transmissions of other users in order to detect

any available unused frequency and use the time or location opportunity to transmit their payloads. The next generation intelligent wireless networks will be centered on this spectrum access technique, where CRs or xG radios can intelligently observe their radio operating environment, detect unused frequency, and dynamically use those holes to communicate while ensuring minimal or no interference with active users (Goldsmith et al., 2009:896 – 897; FCC 2002:2 – 135).

The three spectrum sharing or coexistence techniques aforementioned have distinctive and common features, which are summarized in **Table 2**. In this comparison, underlay and overlay have a common feature of permitting concurrent cognitive and noncognitive transmissions while interweave does not support simultaneous transmission with noncognitive radios or existing users – advocating that a frequency band must exist as spectrum hole before CRs can transmit (opportunistic model). From the interference minimization perspective, interweave spectrum sharing model seems more suitable, since interference with PUs and other active CRs can be avoided by exploiting only spectrum holes.

Underlay	Overlay	Interweave
Channel Side Information: Cognitive (secondary) transmitter knows the channel strengths to noncognitive (primary) receiver (s).	Codebook Side Information: Cognitive nodes know channel gains, codebooks and the messages of the noncognitive users.	Activity Side Information: Cognitive user knows the spectral holes in space, time, or frequency when the noncognitive user is not using these holes.
Cognitive user can transmit simultaneously with noncognitive user as long as interference caused is below an acceptable limit.	Cognitive user can transmit simultaneously with noncognitive user; the interference to noncognitive user can be offset by using part of the cognitive user's power to relay the noncognitive user's message.	Cognitive user transmits simultaneously with a noncognitive user only in the event of a false spectral hole detection.
Cognitive user's transmit power is limited by the interference constraint.	Cognitive user can transmit at any power; the interference to noncognitive users can be offset by relaying the noncognitive user's message.	Cognitive user's transmit power is limited by the range of its spectral hole sensing.

Table 2 Comparison of Underlay, Overlay, and Interweave Cognitive Radio Techniques
(Goldsmith et al., 2009:897)

More so, as illustrated in **Table 2**, each coexistence paradigm relies on different type of information for their operation. Underlay technique requires CRs transmitters to be aware of interference caused to PU_{rx}, overlay technique needs information encoded in the primary users' codebook or messages, while interweave scheme requires the active users' activity information in order to identify spectrum holes. Their power constraints and requirements are also differentiated as shown in the table (Goldsmith et al., 2009:897).

3.   RADIO PROPAGATION AND CHANNEL FADING

The propagation of radio signal through the path between the transmitter and the receiver is susceptible to certain distortive conditions as the signal traverse the air medium. These distortive conditions of the radio channel include but not limited line-of-sight (LOS), obstructions from buildings, mountains, and foliage causing wave propagation phenomena such as reflection, diffraction, and scattering; making radio channel conditions unpredictable. This unpredictable nature of the mobile channel affects the performance of wireless communication systems and makes the analysis of radio channel the most difficult part of radio system design. This channel condition is more severe in urban areas where there is no LOS between the transmitter and the receiver, and the presence of high buildings and monuments cause transmission loss and signal degradation due to reflection and diffraction (Rappaport, 1996: 69 – 70; Rappaport, 2002: 105 – 106; Sklar, 2001:945 – 946).

Over the years, statistical models have been employed in radio engineering to model *wave propagation* in radio channel. Generally, there are two categories of propagation models, namely, *large-scale* models and *small-scale* or *fading models*. Large-scale propagation models are extensively used to predict the mean *signal strength* for radio wave traversing a large distance between the transmitter and the receiver, and they are usually useful in estimating radio coverage area. For example, the *free space model* predicts the received signal power strength as a function of distance, *d,* separating the transmitter and the receiver. Other models in this category are Log-normal shadowing, Longley-Rice model, Okumura model et cetera. In contrast, small-scale models (fading models) deal with the prediction, estimation, and characterization of the rapid fluctuations of the received signal strength over a short distance or time duration (Rappaport, 1996:69 – 70).

For the purpose of this thesis, the focus on propagation models will be *small-scale fading or fading*. Fading will be used throughout this text to refer to small-scale propagation effects. Fading models are primarily used to describe the fluctuation in the magnitude of radio signal's power (amplitude) over a short time duration or distance. Due to reflection, diffraction, and scattering, two or more variations of the same original transmitted signal known as multipath

signals travel through different path and arrive at the receiver at slight different times, causing variation in the amplitude and phase of the received composite signal (Rappaport, 1996:139).

```
                    ┌──────────────────┐
                    │ Fading Channel   │
                    │ Manifestations   │
                    └──────────────────┘
```

Fading Channel Manifestations

Large-scale fading due to motion over large areas — 1

Small-scale fading due to small changes in position — 4

Mean signal-attenuation vs. distance — 2

Variations about the mean — 3

Time spreading of the signal — 5

Time variance of the channel — 6

Time-delay domain Description — 7

Frequency domain Description — 10

Fourier Transform

Time domain Description

Doppler-shift domain Description — 16

Fourier Transform

Duals

Duals

Frequency selective fading — 8

Flat Fading — 9

Frequency selective fading — 11

Flat Fading — 12

Fast Fading — 14

Slow Fading — 15

Fast Fading — 17

Slow Fading — 18

This figure depicts all the fading effects possible in mobile communications, showing the two major classifications, *large scale fading* and *small-scale fading* at the top of the chart.
Figure 16. Fading Channel Manifestations (Sklar 2001: 948)

3.1. RADIO PROPAGATION PHENOMENA

When radio signals propagate from the transmitter and traverse through free space, they are susceptible to reflection, diffraction, and scattering as a result of the effects of ground terrain, atmospheric condition, and physical objects such as buildings, hills, trees etc along the path between the transmitter and the receiver. In most radio technologies such as the mobile communication systems, the realization of line-of-sight LOS path between the transmitter and the receiver is usually unlikely because terrain objects and buildings could have the same or more height than the antenna height, thereby causing reflection, diffraction, and scattering. These phenomena result in *multipath*, a situation whereby the transmitted signal traverses and arrives at the receiver through multiple paths, causing different copies of the original signal to be received

at different time delays with randomly distributed amplitudes and phases. These multiple signals arriving at the receiver are combined to obtain a composite signal, which has varying amplitude and phase depending on the propagation time of the multiple waves and bandwidth of the transmitted signal (Rappaport, 2002: 177; Rappaport, 1996:139).

The three basic radio propagation phenomena contributing to fading are briefly discussed as follows:

1. *Reflection*

    The phenomenon of radio signal reflection occurs when the propagating signal hits the surface of another object (e.g. walls, buildings, trees et cetera), usually object with smooth surface, and the signal is partially transmitted and partially reflected. The dimension of the smooth surface is usually larger than the signal wavelength for reflection to occur (Rappaport, 2002: 177; Rappaport, 1996:139). The impact of reflection on propagating signal bas been discussed in (Jakes, 1974) using the complex analytical result for propagation over a plane earth derived by Norton and simplified by Bullington. This analytical expression considers the direct, reflected, and surface waves to provide a formula that shows the relationship between the transmitted power and the received power as thus:

$$P_r = P_t \left[\lambda/4\pi d \right]^2 g_b g_m |1 + Re^{j\Delta} + (1-R)Ae^{j\Delta} + \ldots|^2 \qquad (3.1)$$

    P_r is the received power, P_t represents the transmitted power, the first, second, and third terms in the series represent the direct wave, reflected wave, and surface wave respectively, where the reflection coefficient R, of the surface depends on the angle of incidence, θ, the polarization of the wave, and the surface characteristics as mathematically expressed in (Jakes, 1974). Also, from equation 3.1, the phase difference between the reflected and the direct paths between the transmitter and the receiver is represented as Δ, as a result of multiple paths created by reflection. **Figure 17** depicts the phenomenon of reflection and multipath effects between the transmitter and the receiver.
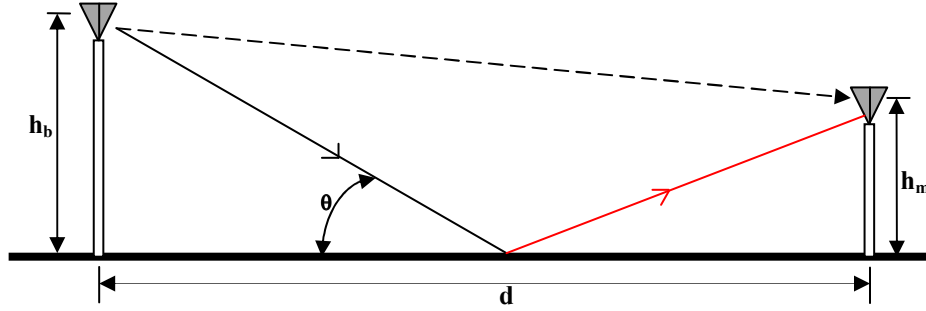
Figure 17.   Propagation paths over a plane earth (Jakes, 1974: 81 – 83)

This phase difference can be derived as shown in equation 3.2 when the heights of the transmitter, h_b (the base) and the receiver, h_m (mobile antenna) are known. Then when the phase difference is estimated, it can be used to evaluate the magnitude of the received power using the complex expression given in equation 3.1.

$$\Delta = \frac{2\pi}{\lambda}\left[\left(\frac{h_b+h_m}{d}\right)^2 + 1\right]^{1/2} - \frac{2\pi d}{\lambda}\left[\left(\frac{h_b-h_m}{d}\right)^2 + 1\right]^{1/2} \tag{3.2}$$

The expression for phase difference shown in equation 3.2 becomes more simplified as in equation 3.3 when the distance, *d,* between the transmitter and the receiver is greater than $5h_b h_m$.

$$\Delta = \frac{4\pi h_b h_m}{\lambda d} \tag{3.3}$$

$$P_r = 4P_o sin^2\left(\frac{2\pi h_b h_m}{\lambda d}\right) \tag{3.4}$$

In free space propagation, the attenuation factor, A is negligible, and if R is defined as -1 equation 4.1 becomes simplified as in 4.4 where P_o is the expected power over free space path (Jakes, 1974: 81 – 83).

2. *Diffraction*

In mobile radio transmission, diffraction occurs when there are obstacles, usually with sharp edges, hindering the existence of LOS between the transmitter and the receiver, such that the secondary waves resulting from the obstructing surface bend around the obstacle. The intensity of diffraction depends on the geometry of the object causing obstruction in the path, and the amplitude, phase, and polarization of the incident wave at

the point of diffraction. This phenomenon has been said to occur frequently in radio propagation as the curved earth surface itself can act as the obstructing object in the path, causing radio signals to propagate around it, beyond the horizon and behind obstructions. In the presence of diffraction effect, the received signal strength decreases as the receiver moves deeper into the region of obstruction (shadowed region). The Fresnel zone geometry has been used to model diffraction and a popular example of such model is the *knife-edge model* (Jakes, 1974:87; Rappaport, 2002:126 - 135). The knife-edge diffraction model as shown in **Figure 18** occurs very often in radio operating environment when the LOS path between the T_x and the R_x is obscured by obstructions such as hills, trees, and tall buildings. Figure 18 shows the worst case scenario when the heights of the transmitter and the receiver are not the same, but the knife-edge diffraction can also be modeled when the heights are equal. The amount of signal attenuation can be estimated by treating the obstructing object as a knife-edge, and assuming that the obstructing object creates a shadow region, the signal strength can be expressed as in equation 5.1 (Jakes, 1974:87; Rappaport, 2002:126 - 135).
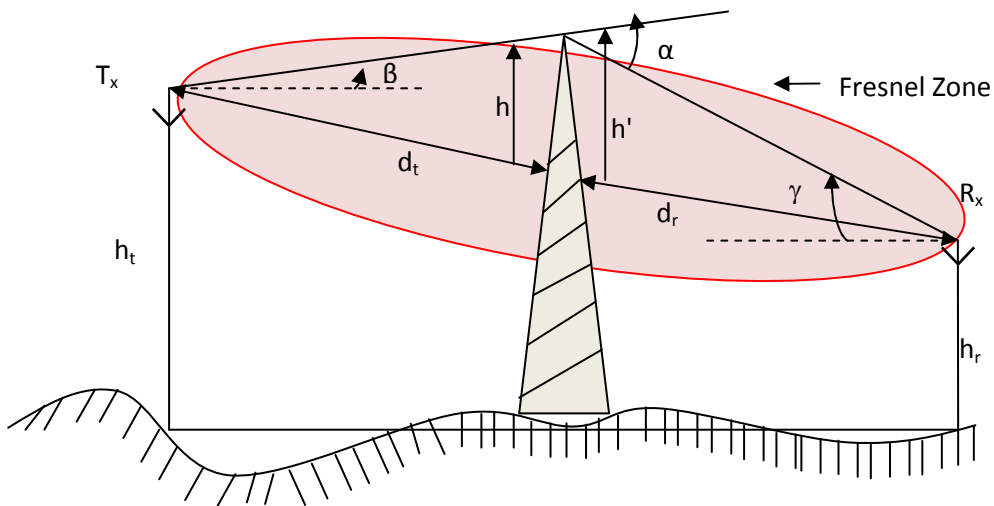
Figure 18.   Knife-edge diffraction geometry with Fresnel zone (Rappaport, 2002:128)

$$\frac{E}{E_0} = Aexp\left(i\theta_p\right) \tag{3.5}$$

Where E0 is the value of electric field at the knife-edge, A is the amplitude, and θp is the phase angle with respect to the direct path. Both θ_p and A can be obtained using the Fresnel integrals as shown in equations 3.6 and 3.7 respectively with the necessary parameters defined in equation 3.8, 3.9 and 3.10 (Jakes, 1974: 87).

$$A = \frac{S + \frac{1}{2}}{\sqrt{2}\sin\left(\theta_p + \frac{\pi}{4}\right)} \tag{3.6}$$

$$\theta_p = tan^{-1}\left(\frac{S + \frac{1}{2}}{C + \frac{1}{2}}\right) - \frac{\pi}{4} \tag{3.7}$$

$$C = \int_0^{h_0} cos\left(\frac{\pi}{2} v^2\right) dv \tag{3.8}$$

$$S = \int_0^{h_0} sin\left(\frac{\pi}{2} v^2\right) dv \tag{3.9}$$

$$h_0 = h\sqrt{\frac{2}{\lambda}\left(\frac{1}{d_t} + \frac{1}{d_r}\right)} \tag{3.10}$$

The above expressions of amplitude and phase angle are more simplified in Rappaport (2002) using the same Figure 18 as the basis. The figure depicts the transmitter and the receiver in free space with obstructing object of height, h between them. The obstructing object is at distance d_t from the transmitter and at distance d_r from the receiver. This causes the propagating radio signal traversing through the obstruction to travel a longer distance than when there is a direct LOS. Therefore, if $h < d_r$, d_t and $h > \lambda$, then the difference between the direct path and the diffracted path, Δ, known as excess path length can be modeled geometrically as:

$$\Delta = \frac{h^2}{2}\frac{(d_t + d_r)}{d_t d_r} \tag{3.11}$$

Hence, the phase difference can be expressed as thus:

$$\phi = \frac{2\pi\Delta}{\lambda} \approx \frac{2\pi}{\lambda} \frac{h^2}{2} \frac{(d_t + d_r)}{d_t d_r} \qquad\qquad (3.12)$$

Basically, in mobile radio communication, diffraction results from the blockage of secondary waves causing a position of the energy to be diffracted around an obstacle; a portion of the signal energy gets blocked in the Fresnel zone while the remaining (unblocked) portion of the signal energy gets transmitted and received at the receiver, hence, the diffraction loss (Rappaport, 2002:126 – 127).

3. *Scattering*

This radio propagation phenomenon is a special case of reflection discussed earlier. It occurs when the radio signal hits a rough surface other than the smooth surface scenario, causing the reflected signal to spread out due to scattering. This could result in the receipt of stronger signal strength at the receiver as opposed to the prediction under reflection and diffraction phenomena. The increase in the received signal strength is as a result of additional radio energy arriving from different directions to the receiver. The intensity of scattering depends on the surface roughness, which can be tested using *Rayleigh criterion*, which defines a critical height (h_c) of surface protuberances for a given angle of incidence θ_i as the parameter, as shown in equation 3.13 (Rappaport, 2002:135 – 136; Sklar, 2001:949).

$$h_c = \frac{\lambda}{8sin\theta_i} \qquad\qquad (3.13)$$

Figure 19 below depicts these three main propagation mechanisms discussed so far, showing a radio signal propagating from a transmitter gets reflected, diffracted, and scattered as it impinges upon the surrounding objects – these objects could be office walls in the case of indoor radio operating environment or buildings, trees, trucks, rocks et cetera in the case of outdoor deployment.
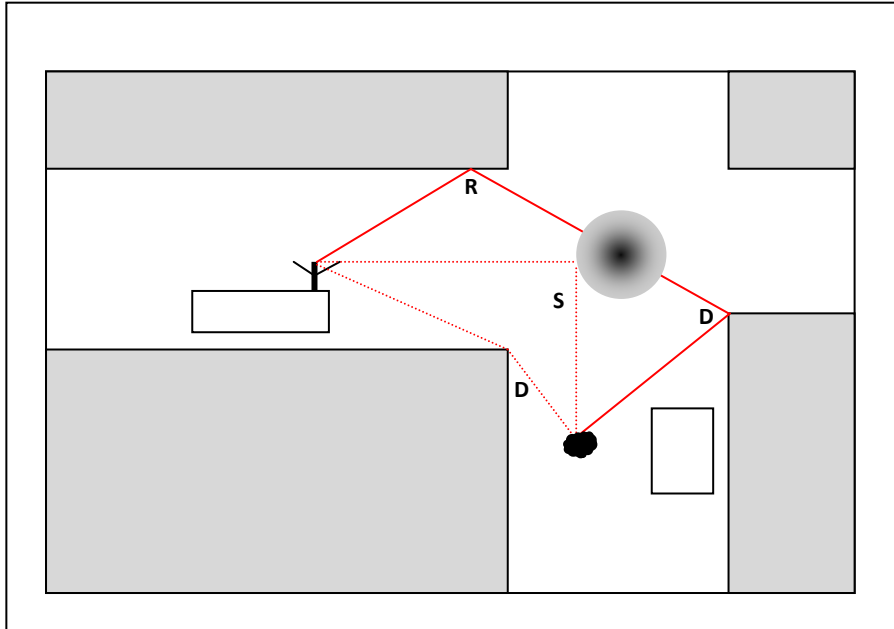
Figure 19. Sketch of three important propagation mechanisms:
reflection (R), scattering (S), diffraction (D)

(Anderson, J. B., et al, 1995: 43)

3.2. SMALL-SCALE FADING AND MULTIPATH EFFECTS

Throughout this text, the term *fading* is used to refer to small-scale fading – a mobile radio propagation phenomenon that describes the rapid fluctuations of the amplitudes, phases, or multipath delays of a radio signal over a short period of time or travel distance given that the large-scale path loss effect is negligible, that is, the effect of large-scale fading is a constant or an assumed value. In other words, fading occurs when two or more replicas of the original signal arriving at the receiver at different times interfere with one another – this interference among the multipath signals results in variation of amplitude and phase of the composite (or resultant) signal formed at the receiver through combining techniques; the magnitude of this variation depends on the distribution of intensity and relative propagation time of the multipath waves and the bandwidth of the transmitted signal. As previously discussed, reflection, diffraction, and scattering phenomena contribute to fading, so multipath in the radio channel creates small-scale fading (Rappaport, 1996:139 – 140; Rappaport, 2002:177 – 178; Sklar, 2001:953 – 957). There are three basic effects associated with small-scale fading and based on these effects, small-scale

fading manifests into two mechanisms, namely, *time-spreading* of the underlying digital pulses within the signal and *time-variant* behavior of the channel due to motion, which are further categorized as illustrated in **Figure 20**. These multipath effects responsible for small-scale fading are as follows:

- Rapid changes in signal strength over small travel distance or time interval

- Random frequency modulation due to varying Doppler shifts on different multipath signals

- Time dispersion (echoes) caused by multipath propagation delays (Rappaport, 2002:177; Sklar, 2001:957).
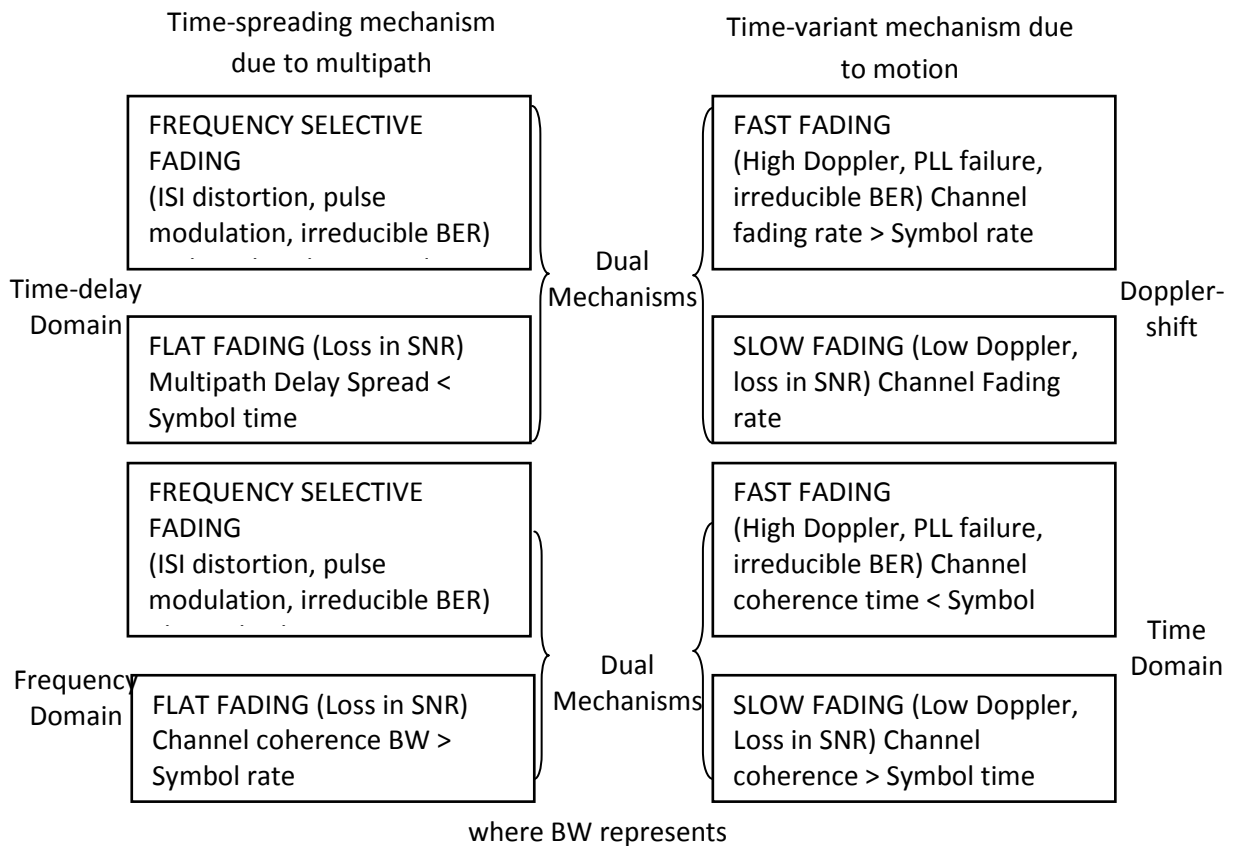
Figure 20. Small-scale fading: Mechanisms, degradation categories, and effects (Sklar, 2001:957)

As the signal travel distance and time interval changes, the signal strength of the received signal changes. The fading phenomenon is more severe in urban radio operating environment where the surrounding structures including growing perennial trees could grow into the Fresnel zone as in the knife-edge illustration in Figure 18. This causes fading because the heights of the mobile

antennas are below that of the surrounding structures, causing blockage of the LOS (Non Line of Sight). On another hand, the blockage of the LOS (NLOS) or obstruction in the Fresnel zone is not only the source of multipath effects causing fading. Reflection caused by these surrounding structures and ground when there is LOS, yields multipath signals traveling through different paths and arriving at the receiving antenna from different directions with different propagation delays. This implies that fading is possible in both LOS and NLOS radio transmission (Rappaport, 2002; 177 – 179).

Receiver's mobility is another factor contributing to multipath propagation. When the mobile receiver is in motion, away from the base station (transmitter), there is a shift in frequency, which is known as the *Doppler Effect* discussed later in this text. This frequency shift is directly proportional to the velocity and the mobile's direction of motion with respect to the direction of arrival of the received multipath wave as illustrated in **Figure 21**. Other factors contributing to small-scale fading are thoroughly discussed in the next section (Rappaport, 2002; 178).

Mathematical and statistical analyses of fading are important tools in understanding and evaluating the performance of radio systems including cognitive radio. The mathematical model of small-scale fading has been given in Sklar (2001) as an analysis with assumption that the antenna remains within a given transmission path in other to assume the large-scale fading, *m(t)* as a constant. In this analysis, the small-scale fading component, $r_o(t)$ is realized as a function of time- variant propagation delay and time-variant multiplicative factor. Assuming that the receiver antenna is in motion and that there are multipath signals created by the antenna's movement, each of these multipath signals is associated with a time-variant propagation delay, $\tau_n(t)$ and a time-variant multiplicative factor $\alpha_n(t)$. The received composite bandpass signal, r(t) follows an equation given in equation 3.14 below. Generally, in signal processing, it is assume that the transmitted signal can be written using complex notation as shown in equation 3.15. Consequently, substituting equation 3.14 in 3.15, the received composite bandpass signal can be written as in equation 7.3 (Sklar, 2001; 953 – 954).

$$r(t) = \sum_n \alpha_n(t)s[t - \tau_n(t)] \qquad (3.14)$$

$$r(t) = Re\{g(t)e^{j2\pi f_c t}\} \tag{3.15}$$

$$r(t) = Re\left(\left\{\sum_n \alpha_n(t)s[t - \tau_n(t)]\right\} e^{j2\pi f_c[t-\tau_n(t)]}\right) \tag{3.16}$$

$$= Re\left(\left\{\sum_n \alpha_n(t)s[t - \tau_n(t)]\right\} e^{j2\pi f_c t}\right) \tag{3.16}$$

$$z(t) = \sum_n \alpha_n(t)e^{-j2\pi f_c \tau_n(t)}g[t - \tau_n(t)] \tag{3.17}$$

3.2.1. Factors Influencing Fading

In practical analysis of radio propagation channel and its performance analysis, *free space propagation model* is insufficient, especially when radio signal propagates through the atmosphere and near the ground. So, signal propagating through the atmosphere could travel through multiple reflective paths and causes fluctuations in received signal strength – this is known as *multipath fading* (Sklar, 2001: 946). In addition to multipath propagation phenomenon, other factors influencing fading in radio propagation channel include but not limited to speed of the mobile, speed of surrounding objects, and the transmission bandwidth of the signal (Rappaport, 2002:178 – 179). Descriptions of these factors are as follows:

1. *Multipath propagation*

   Multipath propagation is undoubtedly the most fundamental factor contributing to fading as most radio operating environment have reflecting and scattering objects present in the radio scene between the transmitter and the receiver. The presence of these physical objects causes scattering of the original signal energy into replicas that traverse different paths to the receiver. This scenario results in the reception of multiple copies of the signal at the receiver with varying phases and amplitudes. Therefore, the reception of different multipath components causes fluctuations or/and degradation in signal strength; hence, fading (Sklar, 2001: 946; Rappaport, 2002:178).

2. *Speed of the mobile*

The fading phenomenon becomes more severe and does not only depend on the multipath effects when the mobile station (usually the receiver) is in motion. This relative motion between the transmitter and the mobile receiver yields a shift in frequency of the multipath components; an effect known as *Doppler shift*, which is discussed in section 3.2.2 of this text. This shift in multipath signals' frequency causes random frequency modulation, and this effect could be positive or negative depending on the mobile's direction of motion as discussed in section 3.2.2 and illustrated in **Figure 20** (Rappaport, 2002; 178).

3. *Speed of surrounding objects*

In some radio propagation paths, the surrounding objects' motion effects contribute and dominate the small-scale fading by inducing time varying Doppler Shift on the multipath components. This effect is negligible and not considered in analyzing the magnitude of small-scale fading, if the speed of the mobile is greater than that of the surrounding objects; else, the effect of the moving surrounding objects has great impact on fading. The *coherence time* is used to determine the degree of surrounding objects' motion by estimating the Doppler shift – what is described as "staticness" of the channel in Rappaport (2002), which refers to the time over which the multipath components are considered coherent (Rappaport, 2002:178 – 179).

4. *The transmission bandwidth of the signal*

The bandwidth of the transmitted signal is an important factor that determines the correlation of the received signal. A distorted signal is received at the receiver if the bandwidth of the transmitted signal is greater than that of the multipath channel. However, this effect does not contribute to fading as the received signal strength is not affected by this distortion (Rappaport, 2002:179). The bandwidth of the channel depends on the coherence bandwidth of the multipath channel, f_o, which according to Sklar (2001) is a statistical measure of the range of frequencies over which the channel passes all spectral components with approximately equal gain and linear phase (Sklar, 2001:960). This implies that the coherence bandwidth is determined by taking the composite of all

frequencies through which the multipath components propagate in the channel. Therefore, if the coherence bandwidth is less than that of the transmitted signal, a distorted signal will be received, but without degradation in signal strength. As culled from Jakes (1974), coherence bandwidth is a useful parameter in analyzing the performance and limitations of different modulation and diversity reception techniques (Jakes, 1974:46). The diversity reception techniques are covered in the chapter four of this text.

3.2.2. The Doppler Shift

The Doppler Effect illustration in **Figure 21** will form the basis of discussion on Doppler shift in this section. As previously discussed, when mobile is in motion farther away from or approaching the base station, the multipath components undergo frequency shift, this shift in received signal frequency is referred to as Doppler Shift (or Doppler Frequency, f_d). **Figure 21** below describes this phenomenon is more practical viewpoint. The figure depicts a moving mobile (truck with radio receiving antenna) at constant velocity v, from point X to point Y of distance, d along the coverage path of a transmitter located at source point, S. The Doppler shift of the multipath components could be estimated if the difference, Δl in path lengths is known. The change in path lengths traveled by the signal from S to the mobile at points X and Y is given as:

$$\Delta l = dcos\theta = v\Delta tcos\theta \tag{3.18}$$

The travel time of the mobile from X to Y is represented as Δt and θ is spatial angle. The change in path lengths causes a change in the phase of the received signal; therefore, the phase difference is estimated as thus:

$$\Delta\phi = \frac{2\pi\Delta l}{\lambda} = \frac{2\pi\Delta t}{\lambda} \, cos\theta \tag{3.19}$$
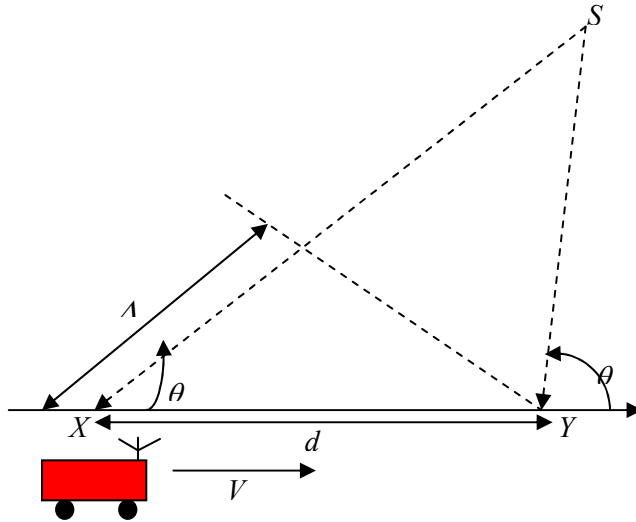
Figure 21. Illustration of Doppler

Hence, Doppler shift equation becomes:

$$f_d = \frac{1}{2\pi}\frac{\Delta\phi}{\Delta t} = \frac{v}{\lambda}\cos\theta \tag{3.20}$$

(Rappaport, 2002: 179 – 180)

From equation 3.20, Doppler shift can also be written as a function of the mobile velocity and the spatial angle. Therefore, it is expected that the frequency of the received signal increases as the mobile moves toward the direction of arrival of the signal – in this case, the Doppler shift is positive and the received frequency, f is the summation of Doppler frequency, f_d and the carrier frequency f_c. On the other hand, the Doppler shift is negative, if the mobile is moving farther away from the direction of arrival of the signal, then the Doppler frequency is subtracted from the carrier frequency to obtain the magnitude of the received frequency, that is, $f = fc – fd$. In Rappaport (2002), a sample calculation of the received frequency is given (Rappaport, 2002:179 – 180).

As culled from Sklar (2001), the effect of Doppler shift is sometimes compensated in radio system design, especially for transmitter synchronization. The techniques used for such design are intended to precorrect the timing and transmission frequency of the signal so that the signal will be received with the expected frequency and at the expected time. This precorrection mechanism works well for a channel with stable radio operating characteristics. The transmitter

performs time precorrection by dividing the distance between itself and the mobile receiver by the speed of light, to estimate the transmission time and adjust the signal transmission time so that the signal arrives at the receiver at the appropriate time. Similarly, for frequency precorrection, the transmitter compensates for the Doppler shift caused by the motion between transmitter and receiver, by adjusting the required transmission frequency as thus:

$$\omega = \left(1 - \frac{V}{c}\right)\omega_0 \qquad\qquad (3.21)$$

where ω is the required transmission radian frequency, c represents the speed of light, V is the relative velocity, which is positive for decreasing transmission distance, and ω_0 is the nominal transmission radian frequency. In principle, the frequency and time precorrections are not straightforward because certain magnitude of errors exists for both techniques (Sklar, 2001:644 – 645).

3.3.   MULTIPATH CHANNEL PARAMETERS

There are several channel parameters that describe the operating characteristics and behavior of multipath channel. These parameters are derived from the power delay profile. This section discusses the four main parameters, namely, coherence time, coherence bandwidth, time dispersion, and Doppler spread. These parameters are used for channel estimation and the design of wireless systems because they quantify channel quality. The average small-scale power delay profile is obtained by taking the average of all instantaneous measurements of power delay profile over a local area (Rappaport, 2002:197).

3.3.1. Power Delay Profile

Power delay profile is used to express in dB scale, the magnitude of the received signal from a multipath channel as function of the delay time, and provides information about the channel parameters such as delay spread. In small-scale channel modeling, the power delay profile is estimated by taking the spatial average of $|h_b(t;\tau)|^2$ in order to obtain a single time-invariant multipath power delay profile, $P(\tau)$. Measurements taken from different radio operating environment give multiple power delay profiles, which contain varieties of different small-scale multipath channel states. The received power delay profile in a local area is obtained by taking

the average over the local scale areas and multiple measurements of $|h_b(t;\tau)|^2$ and it is given by this equation (Rappaport, 2002;184 – 185):

$$P(\tau) = \overline{k|h_b(t;\tau)|^2} \tag{3.22}$$

3.3.2. Time Dispersion

In the design of radio systems and performance analysis of multipath channels, four major time dispersion parameters are used for measuring delay spread of a multipath channel; these are the *mean excess delay, root mean square (rms) delay spread, and excess delay spread*, which are derived from a power delay profile. The mean excess delay and the *rms* delay spread are measured relatively to the first detectable signal received at $\tau_o = 0$ by the receiver. The magnitude of these measurements depends on whether the radio channel is indoor or outdoor, i.e. the characteristics of the measured radio operating environment – usually, for outdoor measurement, rms delay spread is in microseconds while for indoor it ranges in nanoseconds. For example, in Maliatsos et al., (2006), the quantity of power delay spread was measured using seven different radio operating environments, which include open sea, broad and narrow sea passages with or no vegetation, and surroundings that varied from steep rocks to suburban areas (Maliatsos et al., 2006: 2). This implies that the value of any measure multipath channel parameters invariably depends on the radio environment under study. In order to estimate the *mean excess delay*,τ the first moment of the power delay profile is taken as follows:

$$\overline{\tau} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)} \tag{3.23}$$

From this equation, $P(\tau_k)$ is the received power of the delayed *kth* path and τ_k is the *kth* path delay. The *rms delay spread*, σ_τ, is measured by taking the square root of the second moment of the power delay profile, as given in equation 3.23 below:

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\overline{\tau})^2} \tag{3.24}$$

where $\qquad \overline{\tau^2} = \frac{\sum_k P(\tau_k)\tau^2_k}{\sum_k P(\tau_k)}$

The mean excess delay and the rms delay spread equations given in 3.23 and 3.24 do not depend on the power level of $P(\tau)$, but on the relative amplitudes of all multipath components that constitute the power level (Rappaport, 2002:199). As adapted from Rappaport (2002), equations

3.23 and 3.24 can be used to mathematically derive the values of the mean excess delay and the rms delay spread respectively, if the value of the power delay profile is given and the corresponding multipath components delay or path delays are known as shown in Figure 22 below:

Power Delay Profile
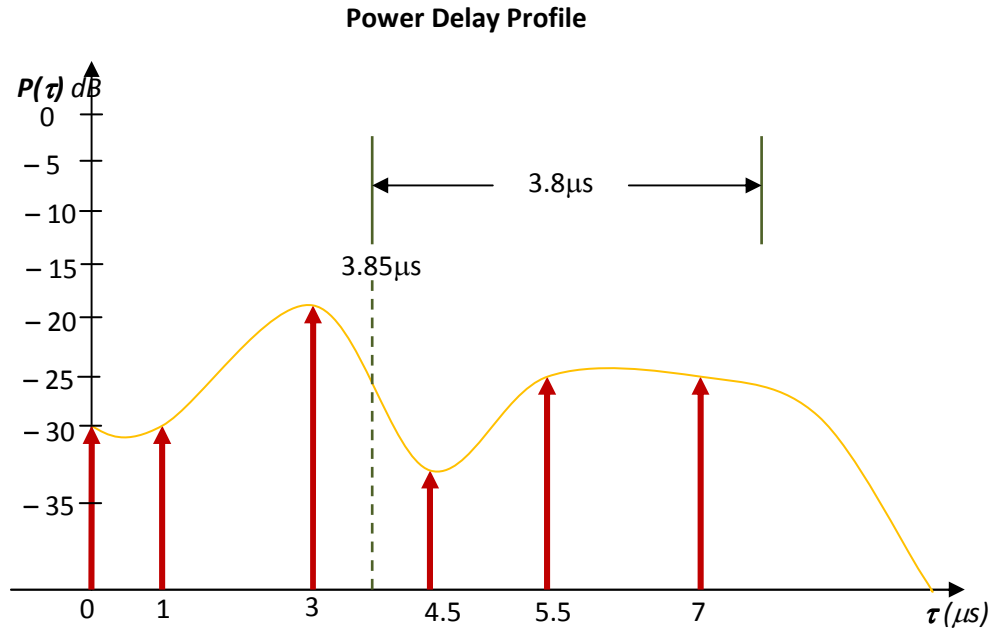
Figure 22.   Example of Root Mean Square (rms) estimation

Firstly, the mean excess delay,τ, which is the first moment of the power delay profile is computed using equation 9.2:

$$\overline{\tau} = \frac{0 + 0.001 + 0.03 + 0.0023 + 0.0174 + 0.02214}{0.001 + 0.001 + 0.01 + 0.00050 + 0.0032 + 0.0032} = 3.85\mu s$$

Estimating the *rms*:

$$\overline{\tau^2} = \frac{0 + 0.001 + 0.09 + 0.0102 + 0.0957 + 0.155}{0.001 + 0.001 + 0.01 + 0.00050 + 0.0032 + 0.0032} = 18.62\mu s^2$$

Finally, applying equation 9.3, the square root of the second moment yields the RMS of 3.8μs below and as indicated on the delay profile of Figure 22 above.

$$\sigma_\tau = \sqrt{18.62 - (3.85)^2} = \sqrt{18.62 - 14.82} = 3.8\mu s$$

The ***maximum excess delay*** (**X dB**) is another important time delay measurement in multipath channel measurement. It is the time delay during which multipath energy falls below the

expected maximum XdB. It defines the extent, which multipath remains above a threshold. In other words, it is the difference between the maximum delay, τ_X at which the multipath component is within certain XdB and the delay τ_o of the first arriving signal, that is, $\tau_X - \tau_o$ (Rappaport, 2002: 199).

3.3.3. Coherence Bandwidth

The study of coherence bandwidth of a multipath channel is as important as the study of delay spreads discussed in previous section because both parameters describe the dispersive nature of the channel in a local area, and are useful in evaluating the performance and shortcomings of different modulation and receiver diversity techniques. According to Jakes (1974), the maximum frequency difference for which the signals are still strongly correlated is called the *coherence bandwidth* of the mobile radio transmission path (Jakes, 1974:46). Similarly, Sklar (2001) defines coherence bandwidth, f_o as a statistical measure of the range of frequencies over which the channel passes all spectral components with approximately equal gain and linear phase. Thus, the coherence bandwidth represents a frequency range over which a signal's frequency components have a strong potential for amplitude correlation (Sklar, 2001:960). In simple words, it is the frequency range at which signal's frequency components, say multipath components have the same amplitude or remain the same. As culled from Rappaport (2002), coherence bandwidth is a relation derived from the rms delay spread, σ_τ and it is the range of frequencies over which the channel is considered "flat;" passing all spectral components with approximately equal gain and linear phase (Rappaport, 2002: 202).

The coherence function has been given in Gustafsson F., et al., (2010) as a normalized version of the cross spectrum, and it is defined as given in equation 3.25. This function can be used to assess the correlation of two signals in frequency domain, and it approaches **1** at frequencies where the two signals become strongly correlated (Gustafsson F., et al., 2010:96).

$$\frac{\Phi_{yx}(\omega)}{\sqrt{\Phi_{yy}(\omega)\Phi_{xx}(\omega)}} \in [0,1] \; for \; all \; \omega \tag{3.25}$$

The correlation function plays an important role in defining or estimating the coherence bandwidth. If the frequency correlation function is above 0.9, then the coherence bandwidth approximately becomes:

$$f_o = {}^1\!/_{50\sigma_\tau} \qquad\qquad (3.26)$$

But if the frequency correlation function is above 0.5, then the coherence bandwidth is defined as thus:

$$f_o = {}^1\!/_{5\sigma_\tau} \qquad\qquad (3.27)$$

(Rappaport, 2002:202; Sklar, 2001:961)

3.3.4. Doppler Spread

While the delay spread and the coherence bandwidth describe the time dispersive characteristics of a channel, coherence time and Doppler spread are parameters that describe the time varying nature of the channel. This time variation is usually caused by mobile's motion or relative motion between transmitter and receiver or movement of an object in the channel, and the time rate of change of the mobile channel leads to spectral broadening. Therefore, Doppler spread, D_s is the measure of this spectral broadening, and it is the range of frequencies over which the received Doppler spectrum is essentially non-zero – that is, frequency dispersion. The spectral line of a pure sinusoidal tone of frequency, fc after being transmitted will have two different components in the range of $f_c - f_d$ to $f_c + f_d$ where f_d is the Doppler spectrum or Doppler shift (Rappaport, 2002:203). **Figure 23** illustrates the frequency of such a sinusoidal tone frequency with its frequency components under Doppler spread.
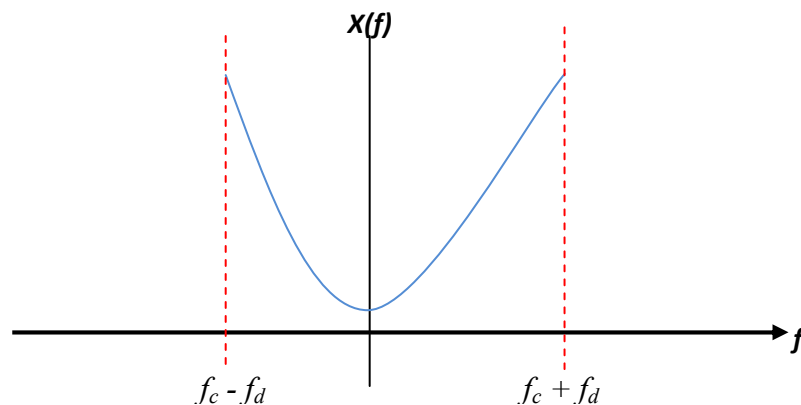
Figure 23.   Sinusoidal tone of frequency under the effect of Doppler spread

3.3.5. Coherence Time

Doppler spread characterizes the time varying nature of frequency dispersion of the channel in the frequency domain while coherence time does so in time domain. In other words, coherence time T_c is the time domain dual of Doppler spread, therefore, it is inversely proportional to Doppler spread as thus (Rappaport, 2002:203 – 204):

$$T_c = {}^1\!/_{D_s} \qquad\qquad (3.28)$$

In Sklar (2001), coherence time is lucidly described using the correlation between a transmitted sinusoidal signal and the impulse response of the channel. **Figure 24** shows the autocorrelation function of the channel's response to the sinusoidal signal. The autocorrelation function, R (Δt) shows the extent (in time) to which the channel's response correlates with the sinusoidal signals transmitted at different time interval t_1 and t_2 ($\Delta t = t_1 - t_2$). Therefore, coherence time, T_c measures the expected time duration over which the channel's response is essentially invariant – correlating with or being similar with the transmitted signal at different time intervals (Sklar, 2001:967).
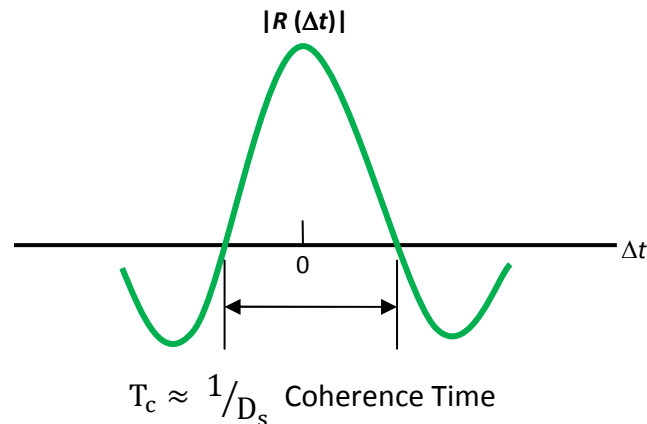
Figure 24. Spaced-time correlation function (Sklar, 2001:959)

3.4. TYPES OF SMALL-SCALE FADING

The aforementioned channel parameters are used to characterize and categorize small-scale fading types as shown earlier in **Figure 20**. In the following subsection, the different types of fading will be discussed, but more focus will be given to *fast fading* and *frequency-selective fading* as they are more important to the main focus of this thesis. The relationship between the signal parameters (such as bandwidth, symbol period etc.) and the channel parameters (e.g.

Doppler spread, Coherence time, etc.) determines the type of fading experienced by signals propagating through multipath radio channel. In other words, channel characteristics in relation with the characteristics of the transmitted signal determine the type of fading. **Figures 20** and **25** show the relationships between the channel parameters based on time delay spread and the signal's properties, and between the channel parameters based on Doppler spread and the signal's properties (Rappaport, 2002:205 – 206).

Small-scale Fading
(Based on multipath timed delay spread)

Flat Fading (Frequency Non-selective)
1. BW of signal < BW of channel
2. Delay spread < Symbol period

Frequency Selective Fading
1. BW of signal > BW of channel
2. Delay spread > Symbol period

Small-scale Fading
(Based on Doppler spread)

Fast Fading
1. High Doppler spread
2. Coherence time < Symbol Period
3. Channel variations faster than baseband signal variations

Slow Fading
1. Low Doppler spread
2. Coherence time > Symbol Period
3. Channel variations slower than baseband signal variations

Figure 25. Types of small-scale Fading (Rappaport, 2002:206)

There are two mechanisms through which small-scale fading manifests, namely, the *time spreading* of the signal and the *time variance* of the channel. From **Figures 20** and **25**, the conditions necessary for time spreading fading and time variance fading to occur are illustrated. Due to *time dispersion* (or time spreading) and *frequency dispersion* phenomena in mobile radio channel, four distinct multipath effects are possible, which are categorized under small-scale fading. Multipath *delay spread* leads to time dispersion and frequency selective fading while *Doppler spread* leads to frequency dispersion and time selective fading. In other words, the two degradation categories experienced by radio signal as a result of the relationship between the maximum delay time, τ_x and the symbol time, T_s (or symbol period) as shown in **Figure 25** are

frequency selective fading and *flat fading* (or frequency non-selective) that occur in multipath channel (Rappaport, 2002:205;Sklar, 2001:959).

3.4.1. Time Spreading Fading

In multipath channel, time dispersion causes the propagating signal components to undergo either *frequency selective fading* or *flat fading*. The intensity of time dispersion and the effects caused by it depends on the relationship between the maximum excess delay time and the symbol period as illustrated in Figure 25. These effects are usually described as fading effects due to multipath time delay spread (Rappaport, 2002:205; Sklar, 2001:958).

3.4.2. Time Variation Fading

The effects of time variation are usually caused by the relative motion between the transmitter and the moving mobile or by the existence of moving objects along the path in the mobile radio channel. This time-varying nature of the channel could cause the propagating signal to experience either *fast fading* or *slow fading* depending on the relationship between the Doppler spread and the propagating signal's characteristics as illustrated in **Figure 25**. These effects are described as multipath effects due to Doppler spread (Rappaport, 2002:208 – 209: Sklar, 2001:966 – 967).

3.4.3. Flat Fading

From figure 25, for flat fading to occur, there are two possible conditions. Firstly, if the bandwidth of the mobile radio channel is greater than the bandwidth of the propagating signal, and the channel bandwidth has a constant gain and linear phase response, then the expected (or received) signal will undergo *flat fading*. Secondly, the channel is said to exhibit flat fading (or frequency non-selective fading) if the delay spread is less than the symbol period (Rappaport, 2002: 205 – 207; Sklar, 2001:959 – 960). Mathematically, a signal undergoes multipath flat fading if $B_s < B_c$ and $T_s > \sigma_\tau$ where B_s is the signal bandwidth, B_c is the channel bandwidth, T_s is the symbol period estimated by taking the inverse of B_s, and σ_τ represents the delay spread as discussed earlier in previous section.

3.4.4. Frequency Selective Fading

Frequency selective fading will occur if the mobile radio channel has a constant gain and linear phase with a bandwidth that is less than the bandwidth of the propagating signal. Additionally, under this same condition for frequency selective fading, the delay spread of the channel impulse response is greater than the symbol period causing the received signal components to undergo fading (or attenuation) and delayed in time, which contribute to signal distortion. In other words, the time dispersion of the transmitted symbols within the mobile channel leads to frequency selective fading, which results in induction of intersymbol interference (ISI) in the mobile radio channel (Rappaport, 2002: 207 – 208). According to Sklar (2001), this time dispersion condition causing ISI distortion occurs when the received multipath components of a symbol extends beyond the actual symbol's time duration, that is, delay spread is greater than the symbol period ($t\sigma_\tau > T_s$). Therefore, frequency selective fading could sometimes be referred to as the channel-induced ISI degradation (Sklar, 2001: 959 – 960). Also, mathematically, a signal undergoes multipath frequency selective fading if $B_s > B_c$ and $T_s < \sigma_\tau$ where B_s is the signal bandwidth, B_c is the channel bandwidth, T_s is the symbol period estimated by taking the inverse of B_s, and σ_τ represents the delay spread.

3.4.5. Fast Fading

Using Figure 25 as reference, fast fading falls under the small-scale fading effects due to Doppler spread. The time variation of the mobile channel causes fast fading under high Doppler spread, when the coherence time of the channel is less than the symbol period of the propagating signal where, due to mobile's motion or the presence of moving objects, the channel experiences rapid changes within the symbol duration – that is, the channel variations are faster than the baseband signal variations. In a simple term, a transmitted signal undergoes *fast fading* when $B_s < B_D$ and $T_s > T_c$ where T_s is the symbol period, T_c represents coherence time of the channel, B_s is the baseband signal bandwidth and B_D is the spectral spread (or Doppler spread) (Rappaport, 2002:208 – 209).

3.4.6. Slow Fading

Invariably, *slow fading* is an opposite effect of *fast fading*. A transmitted signal undergoes slow fading if the rate of changes of the channel impulse response is slower than the baseband signal's

rate of change (or variations). When this occurs, it may be assumed that the mobile radio channel is static or invariant at some time intervals within the symbol duration. And when viewed in the frequency domain, it means that the Doppler spread of the channel is less than the bandwidth of the baseband signal (Rappaport, 2002:209 – 210). This implies a transmitted signal undergoes *fast fading* when $B_s > B_D$ and $T_s < T_c$ where T_s is the symbol period, T_c represents coherence time of the channel, B_s is the baseband signal bandwidth and $\boldsymbol{B_D}$ is the Doppler spread.

3.5.   STATISTICAL DISTRIBUTIONS AND MODELS FOR CHANNEL FADING

In the design and performance analysis of mobile radio systems, fading statistics on certain time intervals have been used to provide useful information related to the statistics of burst errors occurring on fading channels. This statistics of burst errors is used in designing error-correcting codes and optimization of interleaving size (Youssef et al., 1996:1244). In this section, the statistical distributions and stochastic processes related to channel analysis in radio system design will be thoroughly discussed, and the statistical distributions and characteristics of three main flat fading, namely, Nakagami, Rayleigh, and Ricean fading.

3.5.1.   Statistical and Stochastic Processes in Channel Analysis

In fading channel analysis, several models have been proposed to explain the statistical nature of the mobile radio channel. Examples of these models are the Clarke's model and the Jakes' model derived from Clarke's model. In Jakes' model, it is assumed that the fading channel has no LOS, and simulates the received lowpass envelope of a stationary flat fading under isotropic scattering conditions (Rappaport, 2002:214; Patzold and Laue, 1998:712). The accuracy of Jakes' model requires an appropriate stochastic reference model for the received complex lowpass envelope, which follows a complex Gaussian noise process as thus:

$$\mu(t) = \mu_1(t) + j\mu_2(t) \tag{3.29}$$

In equation 11.1 above, $\mu_1(t)$ and $\mu_2(t)$ are uncorrelated zero-mean real Gaussian noise processes with identical variances $\sigma_o^2 = E\{\mu_i^2(t)\}$ (i = 1, 2) where E{.} denotes the statistical average. Therefore, assuming that the analysis is based on an isotropic scattering environment, the autocorrelation function of the stochastic processes $\mu_1(t)$ and $\mu_2(t)$ becomes:

$$r_{\mu\mu}(\tau) = 2\sigma_o^2 J_o(2\pi f_{max}\tau) \tag{3.30a}$$

$$r_{\mu_i \mu_i}(\tau) = \sigma_o^2 J_o(2\pi f_{max}\tau) \tag{3.30b}$$

where $J_o(.)$ is the zero-order Bessel function of the first kind, and f_{max} represents the maximum Doppler frequency. Figure 26 shows schematically, the simulation of Jakes using N weighted low frequency oscillator (Patzold and Laue, 1998:712 – 713). The complete discussion on Clarke's and Jakes' models is outside the scope of this thesis, readers are advised to read Rappaport (2002:214 – 219) and Patzold and Laue, (1998:712 – 718) for more details on Clarke's model and Jakes' model respectively.
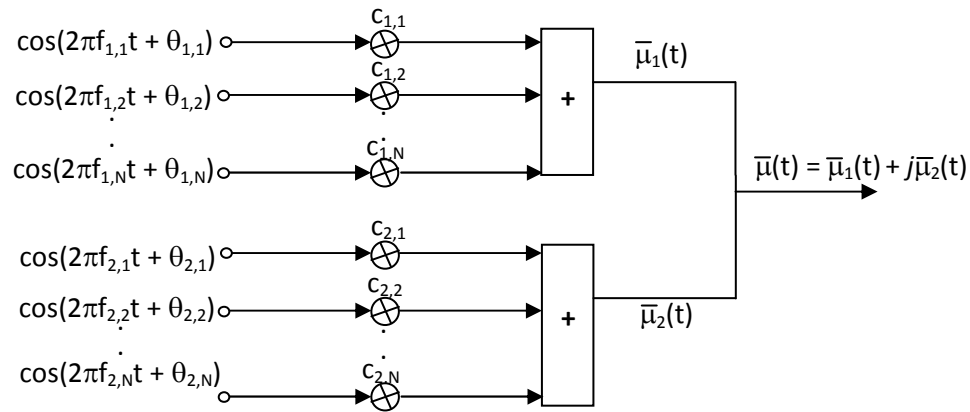
Figure 26.   Jakes' fading simulation with generated complex lowpass envelope using a number of N weighted low-frequency oscillators (Patzold and Laue, 1998:713).

Furthermore, in predictive modeling and analysis of fading channel, certain statistical functions and probability distributions or properties are of great importance. One important function is the characteristic function. The characteristic function of a random variable X is defined for real ω as:

$$\varphi(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} d\, F_X(x) = \int_{-\infty}^{\infty} \cos(\omega x)\, dF_X + i \int_{-\infty}^{\infty} \sin(\omega x)\, d\, F_X(x)$$

It is used to characterize any distribution, and it will always exist because it is always possible to integrate the distribution function over a given interval of τ - representing the inverse Fourier transform of the distribution function. Let the characteristic function, $\varphi(t) = M(it)$, $M(.)$ is the moment generating function of random variable X. Therefore, the characteristic function of sum

of two independent random variables is the product of individual characteristic functions, which determines the distribution function such as the Gaussian distribution. The probability density function (PDF) of random variable X can be obtained by taking the inverse Fourier transform of the characteristic function as shown in equation below:

$$\varphi(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(\omega) e^{-j\omega x} \, d\omega \tag{3.31}$$

Discussion on the basic concepts in probability theory necessary to understand or analyze stochastic process, especially, the deterministic modeling of mobile radio channels, is very essential. It has been shown that distribution function can be obtained from characteristic function of a random variable. The probability distributions that could be used for the analysis and presentation of statistical simulation result include the Gaussian (or normal) distribution, Chi-square distribution, Exponential distribution, and Non-central Chi-square distribution. But the three most important distributions in this context are Nakagami-m, Rayleigh, and Ricean distributions. The first concept is the ***cumulative distribution function*** (CDF), which is define as:

$$F_{\mu} : \mathbb{R} \to [0, 1], x \to F_{\mu}(x) = p(\mu \le x) \tag{3.32}$$

the CDF of the random variable, x and satisfies the following properties:

1. $F_x(-\infty) = 0$ 2. $F_x(\infty) = 1$ and $F_x(x)$ is non-decreasing, that is, $F_x(x_1) \le F_x(x_1)$ if $x_1 \le x_2$.

In other words, $F_x(x) = p(x \le X)$ denotes the probability that the random variable x takes on a value that is less than or equal X. The second important function is the ***probability density function*** (PDF) used to describe a stochastic variable X whose properties are defined in the first two moments, the mean (μ) and the variance (σ^2):

$$E(X) = \mu = \int_x x p_x(x) dx \tag{3.33}$$

$$Var(X) = \sigma^2 = \int_x (x - E(X))^2 p_x(x) dx = E(X^2) - (E(X))^2 \tag{3.34}$$

Another notation is given as:

$$P_x : \mathbb{R} \to 1, x \to P_x(x) = \frac{dF_x(x)}{dx} \tag{3.35}$$

In PDF, the CDF $F\mu(x)$ is differentiated with respect to x. Hence, the PDF satisfies the following properties:

1. $P_x(x) \geq 0 \quad for \ all \ x$

2. $\int_{-\infty}^{\infty} P_x(x) \ dx = 1$

3. $F_x = \int_{-\infty}^{\infty} P_x(x) \ dx$

The **Expected value** *(or mean)* takes the mean average of the random variable x as given by the equation 3.32, where E{.} denotes the expected value operator. If *f(x)* is the function of the random variable x, the expectation value of *f(x)* can be defined as:

$$E\{f(x)\} = \int_{-\infty}^{\infty} f(x) P_x(x) \ dx \tag{3.36}$$

For two random variables x_1 and x_2, the generalization becomes:

$$E\{f(x_1, x_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) P_{x_1, x_2}, (x_1, x_2) \ dx_1 dx_2 \tag{3.37}$$

The **Variance** describes the extent or degree to which approaches its expected value:

$$Var\{x\} = E\{(x - E\{x\})^2\} = E\{x^2\} - (E\{x\})^2 \tag{3.38}$$

The **Moments** describe the Kth moment of the random variable x given as:

$$E\{x^k\} = \int_{-\infty}^{\infty} x^k \ p_x(x) dx, \quad k = 0, 1, ... \tag{3.39}$$

The important probability density functions (PDFs) that are related to this study are defined as follows:

1. **Uniform Distribution**: $f_\theta(x) = \begin{cases} 1/2\pi, & x \in [-\pi, \pi) \\ 0, & else \end{cases}$ where θ is a real-valued random variable and uniformly distributed over the interval $[-\pi, \pi)$.

2. **Gaussian Distribution (or Normal Distribution)**: $P_x(x) = \frac{1}{\sqrt{2\pi}\ \sigma_x} e^{-\frac{(x-m_\mu)^2}{2\sigma_x^2}}, x \in \mathbb{R}$ where x is the real-valued random variable that is Gaussian distributed (or normally distributed), $m_x \in \mathbb{R}$ is the expected value, and $\sigma_x^2 \in (0, \infty)$ is the variance of x: $x \sim N(m_x, \ \sigma_x^2)$.

3. **Multivariate Gaussian Distribution** given as:

   $P_{x_1 x_2 .....x_n}(x_1, x_2 ......, x_n) = \frac{1}{(\sqrt{2\pi})^n \sqrt{det C_x}} e^{-1/2(x-m_x)^T C_x^{-1}(x-m_x)}$ where T is the transpose of a vector (or a matrix), mμ and x are column vectors while $det C_x$ (C^{-1}_x) denotes the determinant (inverse) of the covariance matrix.

(Pätzold, 2002: 11 – 20, Gustafsson et al., 2010: 86 – 87; Ross, 2010: 24 – 55)

Other PDFs useful in fading analysis, namely, Rayleigh, Rice, and Nakagami distributions are part of the discussions on fading types in the next section.

3.5.2. Rayleigh Fading Channel

The previous discussion on fading channel describes that in a multipath fading channel, there are randomly distributed phases, amplitudes, and angles of arrival of multipath components, which are combined at the receiver through the use of an *envelope detector* to obtain resultant signal strength. If there is no LOS path (that is, there exists a typical non-line-of-sight channel) between the mobile receiver and the transmitter, the radio channel is said to exhibit Rayleigh fading because there will be no LOS component and the received signal (envelope) consists of reflection and scattering multipath components at the receiver, and follows Rayleigh distribution (Tang and Zhu, 2003:490; Rappaport, 2002:210 – 212). In (Nakagami, 1958), Rayleigh distribution is one of the three distributions (others are *m*-distribution and Log-normal distribution) that have proven practical applications in the analysis of fading statistics (or observations). And many practical observations have confirmed the use of Rayleigh distribution in both *ionospheric* and *tropospheric* modes of propagation; at least under scattering conditions (Nakagami, 1958:3).

As culled from (Rappaport, 2002; Patzold & Laue, 1998; Nakagami, 1958), the probability density function (PDF) of Rayleigh distribution is defined as:

$$p(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, & (0 \leq r \leq \infty) \\ 0, & r < 0 \end{cases} \tag{3.40}$$

where σ is the *rms* value of the received signal power before envelope detection and σ^2 represents the time-average power of the received signal before envelope detection. Also, the cumulative distribution function (CDF) of Rayleigh showing the probability that the envelope of the received signal does not exceed a specified value of R is given as (Rappaport, 2002: 210 – 211; Nakagami, 1958: 3):

$$P(R) = \Pr(r \leq R) = \int_0^R p(r)dr = 1 - e^{-\frac{R^2}{2\sigma^2}} \tag{3.41}$$

From the statistical distribution viewpoint, let μ_1 and μ_2 be uncorrelated zero-mean real Gaussian noise processes with identical variance σ^2, the absolute value of equation 3.29 under Jakes'

model discussed previously yields a *Rayleigh* process, that is, R(t) = |μ(t)| = |μ₁(t)+ jμ₂ (t)|. In other words, if μ_1 and μ_2 are non-zero independent Gaussian random variables with an identical variance σ^2 and R = $\sqrt{\mu_1^2 + \mu_2^2}$ then, R is the Rayleigh-distributed random variable with PDF and CDF given in equations 3.40 and 3.41 respectively while R^2 obeys chi-square distribution as previously discussed (Patzold and Laue, 1998: 712 – 713). **Figures 27a** and **27b** depict the theoretical and simulated PDF and CDF of Rayleigh fading distribution respectively; see appendix I on how to generate Rayleigh PDF and CDF.
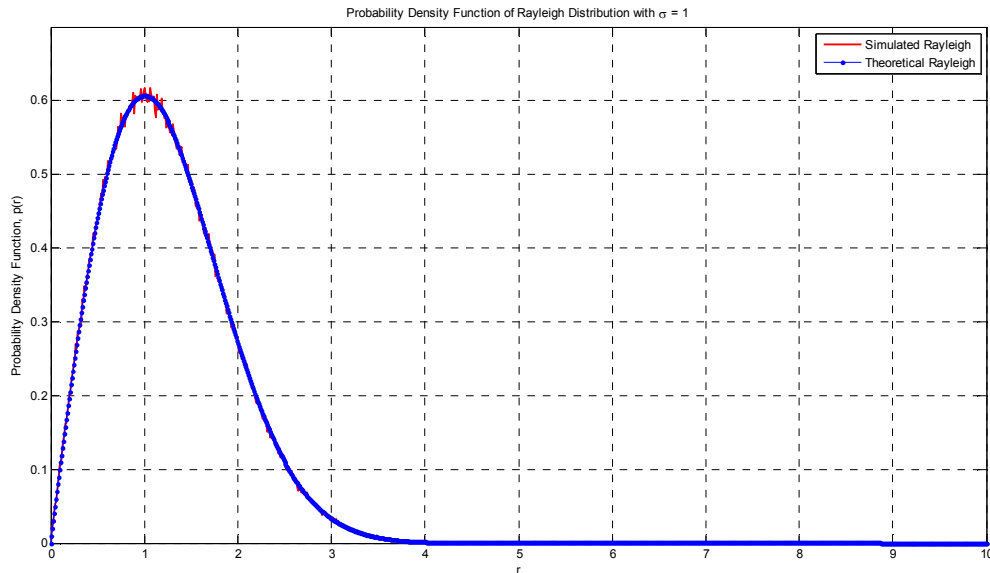
Figure 27a: PDF of Rayleigh fading distribution
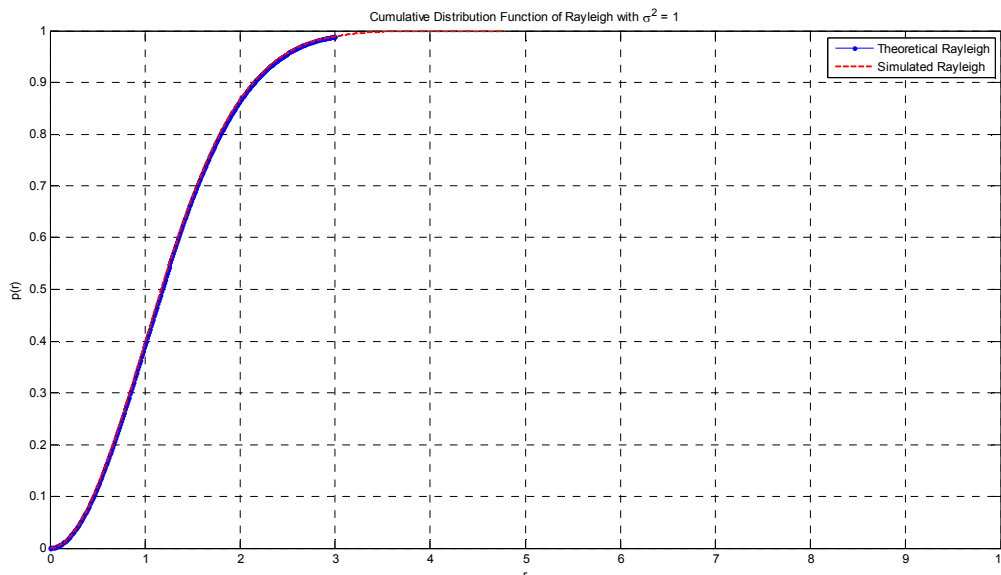
Figure 27b: CDF of Rayleigh fading distribution

3.5.3. Ricean Fading Channel

Recall from previous section that multipath components in mobile radio channel are said to follow Rayleigh distribution if there is no LOS component between the transmitter and the receiver. So, Rayleigh is as a special case of Ricean fading, which occurs when there is a dominant stationary (*non fading*) signal component present among the received multipath components – that is, the small-scale fading envelop distribution follows Ricean distribution if there is a dominant stationary signal component present in it (Tang and Zhu, 2003:490; Rappaport, 2002:212). According to Rappaport (2002), when the fading channel follows Ricean distribution, all the weaker multipath components arriving randomly at the receiver from different angles are superimposed on the stationary dominant signal present in the envelop or composite signal. Rayleigh fading distribution is a special case of Ricean fading because the envelope signal tends to follow Rayleigh distribution if the dominant signal fades away along the path – as the dominant signal becomes weaker the composite signal is received by the receiver as a noise signal with Rayleigh envelope (Rappaport, 2002:212 – 213).

The PDF of Ricean distribution is given as:

$$p(r) = f(x) = \begin{cases} \dfrac{r}{\sigma^2} e^{-\frac{(r^2 + A^2)}{2\sigma^2}} I_0\left(\dfrac{Ar}{\sigma^2}\right), & for (A \geq 0, r \geq 0) \\ 0, & for (r < 0) \end{cases} \qquad (3.42)$$

where A is the peak amplitude of the stationary dominant component, I_o represents the modified Bessel function of the first kind and zero-order while other variables remain as defined under Rayleigh distribution. Also, there is a distinct factor that distinguishes Ricean from Rayleigh, which is the Ricean factor known as parameter K. Parameter K in dB scale is the ratio of the
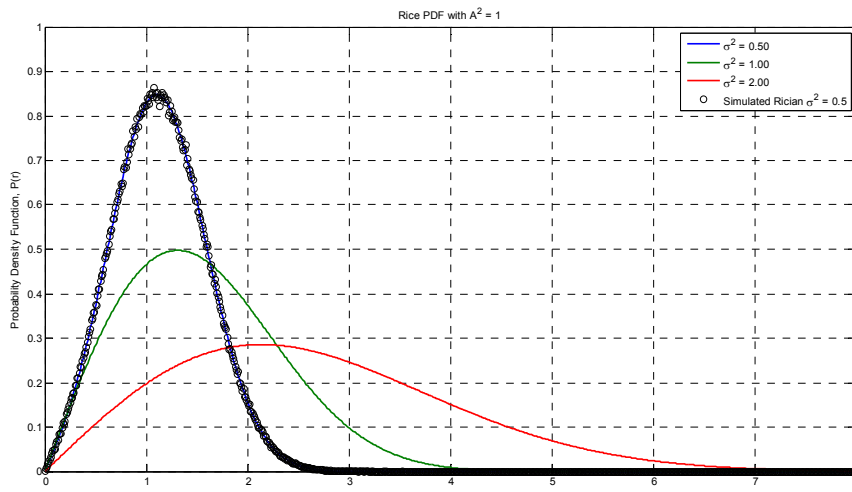
Figure 28: CDF of Ricean fading distribution.

dominant component signal power, A^2 to the variance, σ^2 of the multipath, that is, $K = \frac{A^2}{2\sigma^2}$ in dB

scale it is defined as $K(dB) = 10\log\left(\frac{A^2}{2\sigma^2}\right) dB$ (Rappaport, 2002: 213 − 214). The PDF and

CDF of Ricean fading distribution are shown in Figures 28 and 29 respectively below.

The CDF of Ricean distribution is given as:
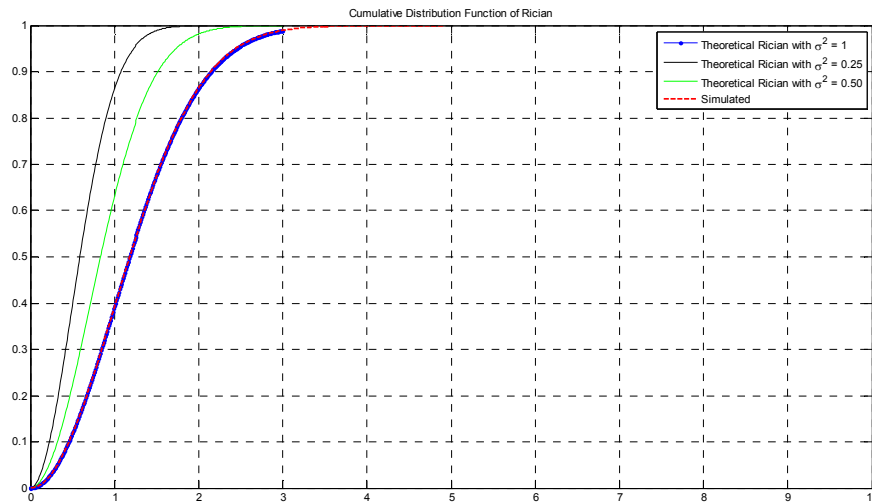
Figure 29: PDF of Ricean fading distribution.

3.5.4. Nakagami Fading Channel

Rayleigh and Ricean fading distributions provide insufficient accuracy in describing long distance fading conditions, or the worst case of fading in mobile channel. As discussed in (Tang & Zhu, 2003), Nakagami fading distribution (or *m*-Distribution) is proposed to compensate the accuracy deficiency in Rayleigh and Ricean models. It was formulated as a parametric gamma distribution-based density function and many researchers have shown that Nakagami distribution is best fit for analyzing fading in mobile communication channel − it could represent various channel fading conditions in radio link (Tang & Zhu, 2003: 490). Having said that Nakagami fading model is more accurate in radio link performance analysis, it includes the Rayleigh and the one sided Gaussian fading models as special cases, and has the flexibility to describe different fading conditions depending on the choice of parameters (Youssef, 1996:1244). The prevalence and universality of Nakagami fading is because the *m*-Distribution uses a parametric gamma distribution-based density function to analyze mobile fading channel rather than

assuming a LOS condition (presence of stationary dominant signal component) as in Ricean fading model. Hence, the PDF of Nakagami distribution is defined as (Tang & Zhu, 2003: 490; Nakagami, 1958: 4):

$$N(r) = \frac{2m^m r^{2m-1}}{\Omega^m \Gamma(m)} e^{-\frac{mr^2}{\Omega}}; \qquad m \geq \frac{1}{2}; \; r \geq 0 \qquad (3.43)$$

where m is the Nakagami parameter, which describes the degree of fading due to scattering and multipath interference process, the average power of the multipath scatter field is given as $\Omega = E[r^2]$ while $\Gamma(m)$ is the gamma function. In a more simplified form, the Nakagami PDF in 3.43 could be written as:

$$f(r) = \frac{2}{\Gamma(k)} \left(\frac{k}{2\sigma^2}\right)^k r^{2k-1} e^{-kr^2/\sigma^2} \qquad r \geq 0; k \geq \frac{1}{2} \qquad (3.44)$$

here, k is the same as the m defining the fading figure as a Gaussian r. v. degrees of freedom. The m parameter in equation 3.42 is further defined in Nakagami (1958) as the inverse of the normalized variance of r^2 as shown in:

$$m = \frac{\left(\overline{r^2}\right)^2}{\left(r^2 - \overline{r^2}\right)^2} \geq \frac{1}{2}, \quad always \qquad (3.45)$$

$$m = \frac{(1+K)^2}{1+2K}, where \; K \geq 0 \qquad (3.46)$$

Nakagami fading model (or distribution) is applicable in modeling tropospheric fading under various conditions. Therefore, it could be described as a generalized form of the Rayleigh model. Again, parameter m represents the depth or the degree of fading as given in 3.45, and Rayleigh exists in Nakagami as a special case when m = 1. When the values of m < 1, the Nakagami fading is said to be more severe, but when m > 1, the fading is less severe and it approaches Ricean fading, and when m → ∞, it approaches Gaussian. The Ricean fading severity parameter K and the Nakagami parameter m could be mapped as in equation 3.46, and **Figure 30** shows the PDF of Nakagami with m parameter equals to 3 (Nakagami, 1958: 4; Bhattacharya, C. et al., 2009: 1; Pajala et al., 2006:82). In a simple term culled from (Yao & Sheikh, 1993), the Nakagami distribution characterizes a generalized channel model, which takes the Rayleigh distribution as a special case and fits experimental data better than Rayleigh, Rician, or lognormal distribution (Yao and Sheikh, 1993:134).
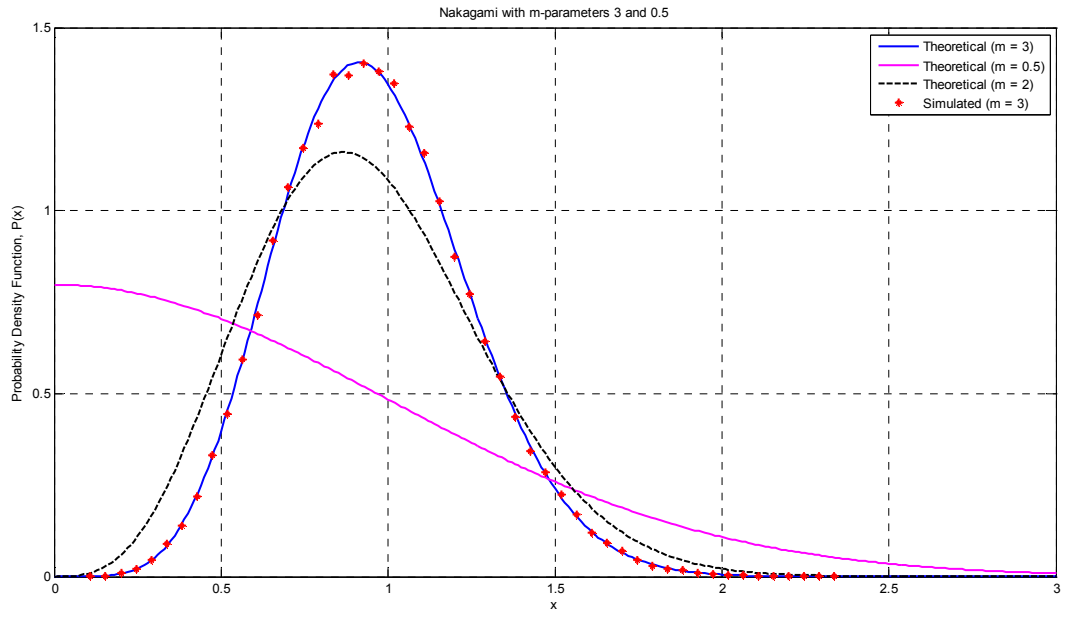
Figure 30:   The PDF of Nakagami Fading Distribution

4. RECEIVER DIVERSITY TECHNIQUES IN RADIO LINK

In most very bad channel condition or radio operating environment, certain signal processing techniques are used to improve the link performance, combat and compensate for the degradation effects due to multipath. The three known categories of techniques used to improve received signal quality and link performance are *channel coding, equalization,* and *diversity*. Each of these techniques can solve a specific channel degradation problem. As discussed earlier, time dispersion multipath effect causes ISI, *equalization* can be used to mitigate the effect of ISI or compensate for it. The ISI situation occurs when the coherence bandwidth is less than the modulation bandwidth, causing modulation pulses to spread in time into adjacent symbols. While equalization compensates for ISI, *channel coding* helps improve the link performance by correcting errors caused by the channel. It is implemented through the addition of redundant data (or extra) bits to the transmitted message for the purpose of correcting any error at the receiver. Examples of channel coding techniques are *turbo code, convolutional code,* and *block code* (Rappaport, 2002: 355 – 356).

The third link performance improvement technique is the *diversity technique*, which is more important in this text. Diversity technique is used to compensate for fading channel impairments, and most diversity techniques are implemented through multiple antenna design mechanisms – placing two or more receiving antenna at the receiver to improve link quality. In other words, diversity technique is used to reduce the depth and duration of the fades reaching the receiver due to motion – combating fading. Diversity techniques have the capability to enhance the link reliability by minimizing the channel fluctuation due to fading and other multipath effects. In its implementation, it is assumed that all the copies of the original signal rarely experience deep fading at the same time – some multipath components will be in deep fade while others will not. Therefore, diversity technique is implemented with multiple or different antennas to receive different versions of the same signal. Unlike equalization and channel coding techniques, diversity technique does not require training overhead or extra bit to operate, therefore, it can improve system performance without adding any overhead or complexity to the radio link. Although, the limitation is that diversity technique is used to compensate small scale fading discussed earlier, but using multiple antennas, independent or uncorrelated signal paths (or

multipath components) are used to improve the received signal-to-noise ratio, SNR (Rappaport, 2002:380 – 381; Sklar, 2001:984 – 985).

As discussed in (Jakes, 1974), diversity technique requires the availability of multiple transmission paths, conveying the same message but having different, uncorrelated, or independent fading statistics, then the mean strengths of the paths is taken and must be approximately the same. In other to ensure that the multiple paths have differing fading statistics, the distance between the receiving antennas must be large enough, which is the concept known as *space diversity reception.* Other forms of diversity include *angle diversity; polarization diversity, frequency diversity, and time diversity* (see Jakes, 1974:310 – 313). Space diversity is easier to implement and does not require additional frequency spectrum, so, it is widely used for compensating channel fading (Jakes, 1974: 310 - 311).

Similarly, as adapted from Rappaport (2002), diversity technique has less complexity because it relies on the random nature of radio propagation by finding independent or uncorrelated signal paths in the channel and there is no need for any transmitter-receiver synchronization as most diversity operations such as selection and combining carried out by the receiver are unknown to the transmitter. The receiver uses multiple antennas to select signals with random strength from different paths and takes the average in order to improve the received SNR (Rappaport, 2002:380). As power control is vastly used to deal with large-scale and path loss degradation, diversity technique is used for combating small-scale fading and it requires adequate spacing of the antennas array so that the incoming multipath components are uncorrelated.

4.1.   RECEIVER DIVERSITY SYSTEM MODEL

This section presents discussion on how diversity techniques can improve the performance of the system and compensate for error caused by multipath fading. As adapted from Sklar (2001), the performance of a radio system with diversity techniques could be analyzed through the estimation of the bit-error-rate (BER), taken as an average, P_B of all multipath fading components and given as:

$$\overline{P_B} = \int_0^\infty P_B(x)p(x)\,dx \tag{4.1}$$

where $P_B(x)$ is the BER probability for a given modulation scheme with a given SNR = x, and x is given as in equation 4.2 below, and $p(x)$ is the PDF of x due to fading.

$$x = \alpha^2 \,{}^{E_b}\!/_{N_0} \tag{4.2}$$

since E_b and N_0 are constants, α is the amplitude variations due to fading effects. Assuming that the incoming fading signals (α) as shown in **Figure 31** below, arriving at each diversity branch are uncorrelated and follows Rayleigh distribution, that is, α^2 and x have a chi-squared distribution, the resultant PDF is in equation 4.3 (Sklar, 2001:985; Jakes, 1974:313 – 314).

$$p(x) = \frac{1}{\Gamma}e^{-\frac{x}{\Gamma}} \quad x \geq 0 \tag{4.3}$$

where the average SNR of all fading elements Γ = SNR = $\overline{\alpha^2}\,{}^{E_b}\!/_{N_0}$ and $x = \alpha^2 \,{}^{E_b}\!/_{N_0}$

Therefore, in an *M- branch* system, each branch $I = 1, 2\ldots M$ has an instantaneous SNR = γ_i and given that the average SNR, Γ is the same for all branches. The following section provides more detailed information on all the diversity techniques and the derivation of each technique.
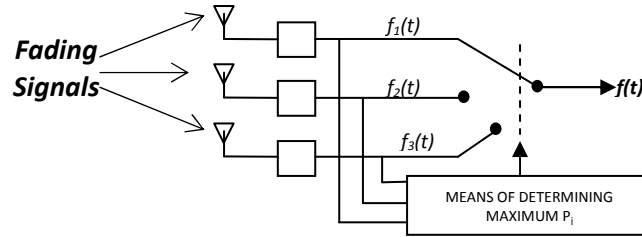
Figure 31. Three-Channel Selection Diversity, where f_j represents incoming multipath component (Brennan, 2003: 336).

Therefore, assuming an RF system with a single user, the received signal is the sum of the transmitted signal, x(t), channel impulse, h, and the noise, n that is, $y = hx(t) + n$. The average power in the composite or noisy signal, y received by an *M-branch* receiver over a symbol period, τ is given as:

$$p = \frac{1}{\tau}\int_0^\tau |x(t)|^2 |h(t)|^2\,dt = h(t) \tag{4.4}$$

assuming that the channel is slow fading and *x(t)* has a unit power. Then the expected variations in the signal power, $\sigma^2 = E\{|n(t)|^2\}$, the instantaneous SNR, γ_i at the ith branch of the *M-branch system* becomes:

$$\gamma_i = \frac{|h_i|^2}{\sigma^2} \tag{4.5}$$

For flat Rayleigh fading where h_i is uniformly distributed over $[0, 2\pi)$ and $|h_i|$ has a Rayleigh PDF as in equation 13.1 (described as density function in Jakes, 1974:313) and could be written as:

$$|h_i| \sim \frac{|h_i|}{\sigma^2} e^{-\frac{|h_i|^2}{2\sigma^2}} \tag{4.6}$$

then the exponentially distributed instantaneous SNR at each receiver become:

$$\gamma_i = \frac{1}{\Gamma} e^{-\frac{\gamma_i}{\Gamma}} \tag{4.7}$$

where Γ is the average receivable SNR at each *M branch i = 1, 2...M.* That is, SNR at each antenna in a diversity system with *M* array of antennas. In the analysis of diversity systems, two figures of merits are important, which are outage probability (or probability of outage) and the bit − error − rate (BER). The probability of outage in this context is the probability that the received SNR, γ at the receiver is less than a predefined threshold, γ_o. The BER is modeled by integrating equation 4.7 and it is given as:

$$P_o = P(\gamma < \gamma_o) = \int_0^{\gamma_o} \frac{1}{\Gamma} e^{-\frac{\gamma}{\Gamma}} \, d\gamma = 1 - e^{-\gamma_o/\Gamma} \tag{4.8}.$$

In general, the outage probability P_o and the BER are suitable figures in fading channel analysis, and decrease in error rate depends on the variance of the SNR due to the randomness of multipath components. For instance, the BER in a Binary Phase Shift Keying (BPSK) system is defined as:

$$erfc\left(\sqrt{2\gamma}\right) = Q\left(\frac{|h|}{\sigma}\right), where \; Q(x) = \left(\frac{1}{\sqrt{2\pi}}\right) \int_x^\infty e^{-t^2/2} \, dt \tag{4.9}$$

then over Rayleigh fading, the BER becomes (Sklar, 2001:984 − 985; Rappaport, 2002: 381):

$$BER = \int_0^\infty \frac{|h_i|}{\sigma^2} e^{-\frac{|h_i|}{2\sigma^2}} Q\left(\frac{|h|}{\sigma}\right) \, d(|h_i|) \tag{4.10}$$

4.2.   RECEIVED SIGNALS COMBINING TECHNIQUES

The *M* antennas receive different uncorrelated components and pass them to the *M-branch* combiner, which uses certain statistical signal processing algorithm (*combining*) to produce a strong SNR. There are three common combining techniques uncovered in most literatures, which are *maximal ratio combining* (MRC), *selection combining,* and *equal gain combining* (EGC). Another type of combining technique discussed in Brennan (2003) is *scanning diversity*, which Sklar (2001) and Rappaport (2002) referred to as *feedback technique* (Brennan, 2003: 335 – 340; Sklar, 2001: 984 – 986; Rappaport, 2002: 385 – 387), but only the common three techniques are discussed in this section. Since some of these combining techniques are derived from the approximation or modification of the other, it is ideal to discuss them in the following order.

4.2.1.   SELECTION COMBINING

Selection combining technique is the easiest technique to implement, but it does not provide optimal SNR because (*N – 1*) elements of the *M-branch* array are ignored, that is, one out of all the received signals is used while discarding others. As previous shown in **Figure 31**, a comparator compares $f_1, f_2,$ *and* f_3 to determine the *M* receiver with the maximum baseband SNR and this *M* receiver with the best SNR is connected to the demodulator. As detailed in Brennan (2003), this form of combining technique does not allow other noisy signals (or multipath components) to contribute to the overall SNR of *f(t)* because only the element with the maximum SNR is chosen. Therefore, such system could be modeled as follows: (Rappaport, 2002:386; Sklar, 2001:986; Brennan, 2003: 336).

$$f_j = f(x) = \begin{cases} 1, & \gamma_j = max\{\gamma_n\} \\ 0, & otherwise \end{cases} \tag{4.11}$$

given that $\gamma_j \geq \gamma_n$ for j = 1, 2 …n.

For selection diversity analysis, we can assume *M* independent Rayleigh fading channels, and each fading channel represents each element in the *M* diversity branch. Each ith branch has the same average SNR as given by equation 4.4. Recall that in Rayleigh fading channels, the fading amplitude, α has a Rayleigh distribution such that the fading power, α^2 and the instantaneous SNR, have a chi-squared distribution with two degrees of freedom, then the PDF expressed in

terms of the SNR is given in equation 4.3 and when $\overline{\alpha^2} = 1$, Γ becomes the average E_b/N_o for the fading channel (Rappaport, 2002:381; Rappaport, 2002:341).

So, each channel or ith branch with instantaneous SNR, γ_i have the same average SNR Γ and that $\overline{\alpha^2} = 1$, the Rayleigh PDF of γ_i is expressed as:

$$p(\gamma_i) = \frac{1}{\Gamma} e^{-\frac{\gamma_i}{\Gamma}}, \qquad \gamma_i \geq 0 \tag{4.12}$$

again, where Γ represents the average SNR of each branch and γ_i is the instantaneous SNR at each branch, say the receiver of f_1, f_2, and f_3 in **Figure 31**. Hence, the probability that a single i^{th} branch has an instantaneous SNR less than the predefined threshold γ_o is given as:

$$P_r[\gamma_i \leq \gamma_o] = \int_0^{\gamma_o} p(\gamma_i)\, d\gamma_i = \int_0^{\gamma_o} \frac{1}{\Gamma} e^{-\frac{\gamma_i}{\Gamma}}\, d\gamma_i = 1 - e^{-\frac{\gamma_o}{\Gamma}} \tag{4.13}$$

The received signal on each receiver branch (or antenna) is independent. Therefore, the joint probability is the product of all probabilities from all the *M branches*. From equation 4.13 above, the probability that the received SNR at all independent diversity branches (i = 1, 2…*M*) is less than the SNR threshold becomes:

$$P_r[\gamma_1, \gamma_2 \cdots \gamma_M \leq \gamma_o] = \left(1 - e^{-\frac{\gamma_o}{\Gamma}}\right)^M \tag{4.14}$$

Therefore, when *selection diversity* technique is used, the probability of the received SNR at an ith branch exceeding the threshold γ_o is given as:

$$P_r[\gamma_i > \gamma_o] = 1 - \left(1 - e^{-\frac{\gamma_o}{\Gamma}}\right)^M = P_M(\gamma_o) \tag{4.15}$$

where $P_M(\gamma_o)$ represents the distribution of the signal with the best SNR selected from the M branches, that is, the probability that the received SNR exceeds a threshold when *selection diversity* is used, and the ith receiver with this strongest SNR is connected to the demodulator. **Figure 32** and **Figure 33** show the plots of $P_M(\gamma_o)$ for diversity system with twenty (20) M

branches and BER for BPSK with selection diversity over Rayleigh fading respectively (Sklar,
2001:985; Rappaport,
2002:381; Jakes,
1974:313 − 316).
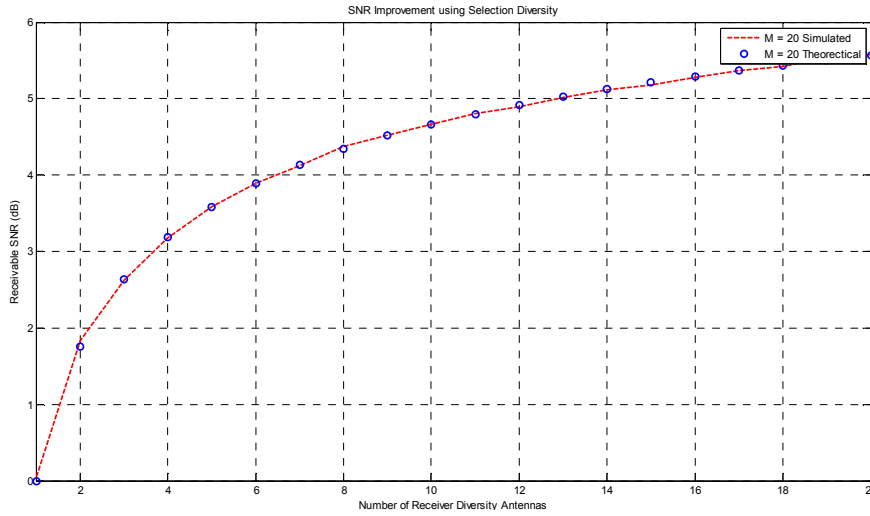
Figure 32. SNR Improvement with Selection Diversity
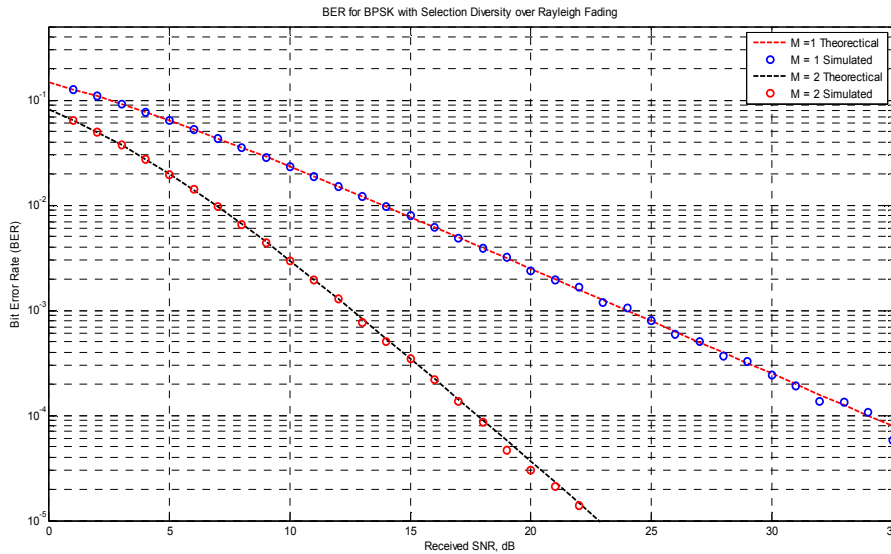

Figure 33. BER for BPSK with Selection Diversity over Rayleigh Fading

4.2.2. MAXIMAL RATIO COMBINING

Recall that this thesis investigated the capacity of CR spectrum sharing with maximal ratio
combining (MRC) receivers over Nakagami fading. Therefore, the discussion and the derivation
of MRC are more important in this context. In maximal ratio combining, the optimal solution

deficiency in selection diversity is improved by obtaining the weighted value of all signals' SNR received at all the *M branches* in order to maximize the resultant SNR at the demodulator; as depicted in **Figure 34**. Unlike selection diversity that selects the signal with strongest SNR, MRC considers all incoming fading components (or signals) in estimating the final received SNR (Rappaport, 2002:387; Sklar, 2001:986; Jakes, 1974:316 - 317).
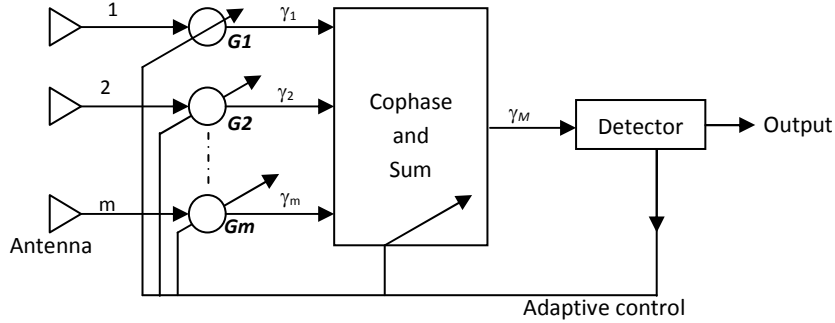
Figure 34. Block Diagram of Maximal Ratio Combining diversity technique (Rappaport, 2002:387)

In other words, MRC yields an average SNR equal to the sum of the individual average SNRs and is given by equation 4.16 with an assumption that each branch has the same average SNRγ_i. Rather than comparing all the received SNRs to a threshold SNR, MRC technique produces an acceptable output SNR (average) even when none of the received SNRs is acceptable (Sklar, 2001:986).

$$\overline{\gamma_M} = \sum_{i=1}^{M} \gamma_i = \sum_{i=1}^{M} \Gamma = M\Gamma \tag{4.16}$$

Consider a system whose composite signal with fluctuating amplitude could be described as the weighted sum (general linear combination) of f_j components given in equation 4.17 below - MRC yields the maximum acceptable SNR. So, let P denote the local power ratio of signal $f(t)$, then MRC yields $P = \sum_{j=1}^{N} p_j$, which is the maximum power obtainable from the linear combination of $f(t)$ in equation 4.17. This solution provides an improvement over selection diversity that considers only one of the terms in the sum $\sum p_j$ (Brennan, 2003:336).

$$f(t) = a_1 f_1(t) + \cdots + a_N f_N(t) = \sum_{j=1}^{N} a_j f_j(t) \tag{4.17}$$

In order to produce largest SNR at the receiver, MRC technique uses each of the M branches in a cophased and weighted manner. That is, the M signals are weighted proportionately to their SNRs and then summed – the individual signals must be cophased before combining, and this is usually achieved by *cophasing* the antennas at a certain distance apart (usually about ½ wavelength or more) and feeding the antennas in-phase (Sklar, 2001:986; Jakes, 1974:316; Rappaport, 2002:387).

Provided the individual signals have been cophased, the envelope of the composite signal is given in equation 4.18 where r_i represents the voltage signals received at each M diversity branch. Before combining, these voltage signals are cophased to provide coherent voltage addition and weighted to provide optimal SNR. In the combined signal envelope of equation 17.3, G_i is the gain at each branch (Rappaport, 2002:384 – 385; Jakes 1974:318 – 319).

$$r_M = \sum_{i=1}^{M} G_i r_i \tag{4.18}$$

Also, MRC improvement takes the total noise power, N_T at the detector into consideration. So, assuming that all M branches have equal average noise power, N, that is, $\overline{n_i^2} = N$ for all i branches, the total noise power, N_T at the detector is given by equation 4.19 as the weighted sum of the noise in each branch (Rappaport, 2002:384 – 385; Jakes 1974:318 – 319).

$$N_T = N \sum_{i=1}^{M} G_i^2 \tag{4.19}$$

From equations 4.18 and 4.19 above, the resultant SNR at the detector is:

$$\gamma_M = {r^2 M}/{2N_T} \tag{4.20}$$

This magnitude of SNR, γ_M at the detector is maximized when $G_i = {r_i}/{N}$, and it can be derived using Chebychev's inequality as thus:

$$\gamma_M = \frac{1}{2} \frac{\Sigma \left({r_i^2}/{N} \right)^2}{\left({r_i^2}/{N^2} \right)} = \frac{1}{2} \sum_{i=1}^{M} \frac{r_i^2}{N} = \sum_{i=1}^{M} \gamma_i \tag{4.21}$$

The derivation of maximum SNR in 4.20 using Chebychev's inequality implies that the SNR produced by MRC at the detector is simply the sum of the SNRs at each branch, $\gamma_i = \frac{r_i^2}{2N}$. In order to model MRC using the received signal envelope in fading channel, we generate two independent Gaussian random variables x and y having zero mean and equal variance σ^2, then

$$\gamma_i = \frac{r_i^2}{2N} = \frac{(x^2 + y^2)}{2N} \tag{4.22}$$

Then γ_M is a chi-squared distribution of 2M Gaussian random variables with variance $\sigma^2/_{(2N)} = \frac{\Gamma}{2}$, such that the PDF of γ_M becomes:

$$p(\gamma_M) = \frac{\gamma_M^{M-1}}{\Gamma^M (M-1)!} \quad for \; \gamma_M \geq 0 \tag{4.23}$$

Therefore, from equation 4.23, we can obtain the probability distribution for MRC by deriving the probability that γ_M is less than a predefined or expected SNR threshold γ_o :

$$P_r\{\gamma_M \leq \gamma\} = \int_0^\gamma p(\gamma_M)\, d\gamma_M = 1 - e^{-\gamma/_\Gamma} \sum_{k=1}^M \frac{(\gamma/_\Gamma)^{k-1}}{(k-1)!} \tag{4.24}$$

In essence, the average SNR, $\overline{\gamma_M}$ in MRC is obtained as the sum of the individual average SNR from each branch in the diversity system as given in equation 4.16. **Figures 35** and **36** respectively show the plots of SNR improvement with MRC given M number of antennas equal to 15 and the BER in Rayleigh fading with BPSK modulation scheme.
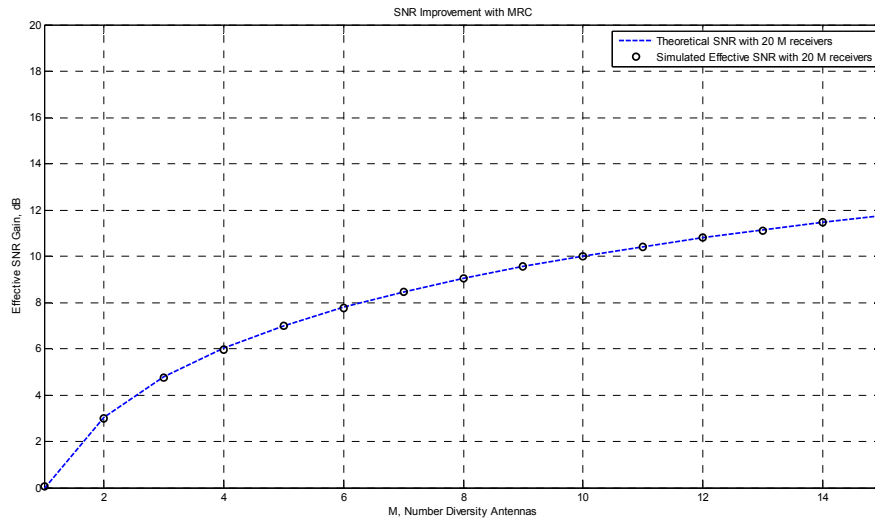
Figure 35. SNR Improvement with MRC, 15 Receiver Antennas
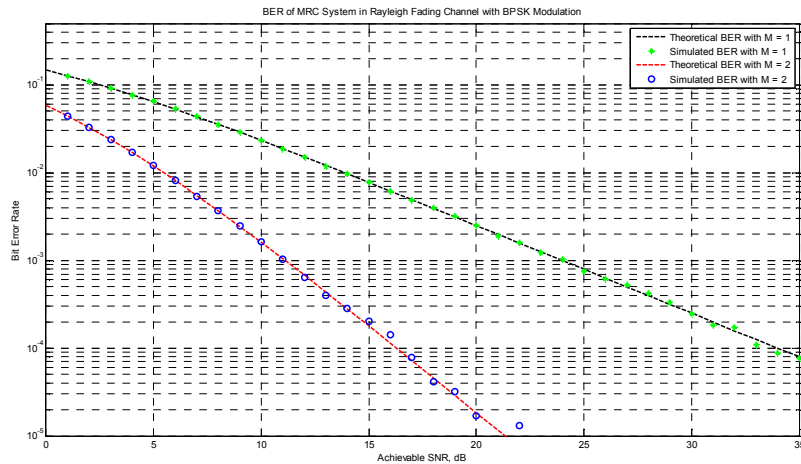
Figure 36. BER of MRC System in Rayleigh Fading with BPSK Modulation

4.2.3. EQUAL GAIN COMBINING

Maximal ratio combining requires the weights from all branches to vary in order to yield optimal SNR. Therefore, because it is not convenient to obtain the variable weighting, Equal Gain Combining (EGC) diversity is used to overcome this problem by setting all the branch weights to unity and cophased the signals from each branch to provide equal gain combining. The derivation is similar to MRC except that the weights are set to unity. Reader is hereby referred to Jakes (1974:319 – 321) and Brennan, (2003:339 – 340) for the derivation of EGC (Rappaport, 2002: 387; Sklar, 2001:986; Jakes, 1974:319). In Brennan (2003), EGC is described as the simplest possible linear diversity technique with an assumption that all channels have exactly the same gain, such that $a_j = 1$ in equation 4.17. Therefore, in simple words, EGC technique simply adds all the noisy signals $f_j(t)$ together (Brennan, 2003:337).

4.3. COMPARING THE DIVERSITY TECHNIQUES

The simple implementation and less design complexity of selection diversity make it preferable to MRC and EGC. However, it does not provide optimal solution (SNR Improvement) as MRC does, which makes MRC a better option for diversity system design – MRC has an advantage of producing an output with an acceptable SNR even when none of the individual signals are acceptable (such as in worst channel condition), so, modern DSP techniques and receivers have MRC implemented for diversity combining. The downside of MRC is the required variable

weighting capability, which is not always convenient or desirable to provide. Therefore, EGC provides a solution to this deficiency by setting the gains from all the *M* receivers equal to a constant value of unity. Table 3 below summarizes the differences between Selection Diversity, MRC, and EGC (Rappaport, 2002: 385 – 387; Brennan, 336 – 337; Jakes, 1974:313 – 321; Goldsmith, 2005: 207 – 217).

Table 3:   Comparing Selection Diversity, MRC, and EGC Diversity Technique

Selection Diversity	Maximal Ratio Combining	Equal Gain Combining
Simple implementation and less design complexity.	Introduces more complexity in diversity system design than selection diversity and EGC.	The complexity in EGC is more than that of selection diversity, but minimal compared to MRC. The tradeoff for the reduced complexity in EGC is 1dB of power penalty.
Does not require knowledge of the time-varying SNR on each branch since only the branch with the highest SNR is selected and connected to the output.	Requires knowledge of each time-varying SNR magnitude on all *M* branches, and this could introduce system processing overhead or design complexity. Also, the time-varying SNR could sometimes be difficult to measure	Does not require knowledge of the time-varying SNR on each branch, but the signals on each branch are cophased and combined with equal weighing.
It is not optimal because only the ith branch with maximum SNR is connected to the decoder while the SNR values of other branches are discarded.	Provides optimal solution – all the received SNRs are considered in the weighted sum operation, hence, final received SNR is maximized.	The solution and performance is somehow close to that of MRC.
Does not provide optimum solution, hence, capacity improvement is relatively minimal compare to MRC.	MRC provides optimum solution; hence, the capacity improvement is relatively maximal to other diversity techniques.	Although, performance is close to that of MRC but capacity improvement is relatively minimal compare to MRC.

4.4.   COGNITIVE RADIO WITH MAXIMAL RATIO COMBINING OVER NAKAGAMI-*M*

In (Li, 2012), the performance analysis of maximal ratio combining technique for cognitive radio was investigated, and the result shows that MRC can provide full diversity order and the capacity scales as a logarithmic function of the number of CR receiver antennas (Li, 2012: 849). In CR system implementation with MRC, the primary user (PU) and the secondary user (SU) have one single antenna while the SU_{rx} has multiple antennas for MRC diversity technique as shown in **Figure 37**. As previously discussed, one of the major concerns in cognitive radio designs and

implementation is the avoidance of harmful interference with the primary user (PU). Therefore, in CR with MRC, the transmit power of the SU_{tx} must satisfy the interference power constraints Q.
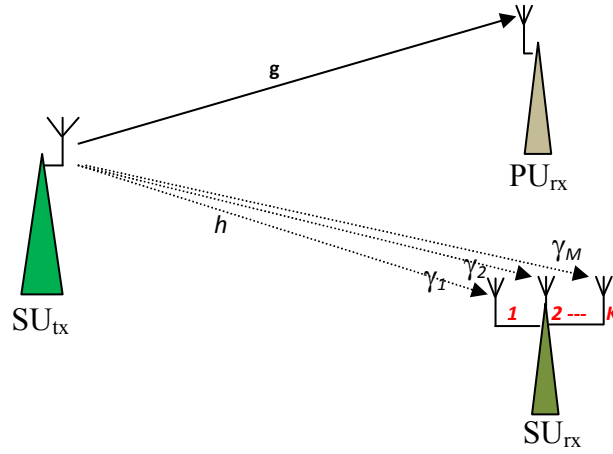
Figure 37. Cognitive Radio System with MRC at Secondary Receiver

Using the diversity system model in **Figure 37** above and assuming that the channel fading is Nakagami, the channel gain is i.i.d and has probability characteristics of Nakagami PDF given previously in equation 3.43, therefore, the PDF of the channel gains g and h for $SU_{tx} \rightarrow PU_{rx}$ and $SU_{tx} \rightarrow SU_{rx}$ respectively, are given in equation 4.25 and 4.26 respectively with an assumption that the average channel gain on both channels is one (Yao & Asrar, 1993:134; Alouini & Andrea, 1997:359; Duan et. al, 2010:2):

$$P_h(h) = \frac{m^m h^{m-1}}{\Gamma(m)} e^{-mh} \tag{4.25}$$

$$P_g(g) = \frac{m^m g^{m-1}}{\Gamma(m)} e^{-mg} \tag{4.26}$$

With MRC at the secondary receiver (SU_{rx}) as shown above, there are K antenna elements, $K \in (1, 2, 3 \ldots K)$, then the PDF of the SNR at the combiner's output becomes (Alouini & Andrea, 1997:359):

$$P_{hMRC}(h) = \frac{m^{Km} h^{Km-1}}{\Gamma(Km)} e^{-mh} \tag{4.27}$$

Furthermore, considering the transmit power constraints and assuming that the Channel Side Information (CSI) is known to both transmitter and receiver, the transmit power, $P_t(\gamma)$ varies with SNR $\gamma[i]$ and it is subject to the interference power constraints Q as thus:

$$\int_\gamma P_t(\gamma) P_{\gamma MRC}(\gamma) \quad d\gamma \leq Q \tag{4.28}$$

5.   SYSTEM MODEL AND NUMERICAL RESULTS

This chapter deals with the system model, the simulation results, and analysis of the results. It is assumed that the channel side information (CSI) is available or known to both the transmitter and the receiver as shown in **Figure 38** below and the MRC antenna configuration in the CR system of **Figure 37**. Transmit power control is implemented in the transmitter and the receiver has channel estimator to keep track of the CSI.
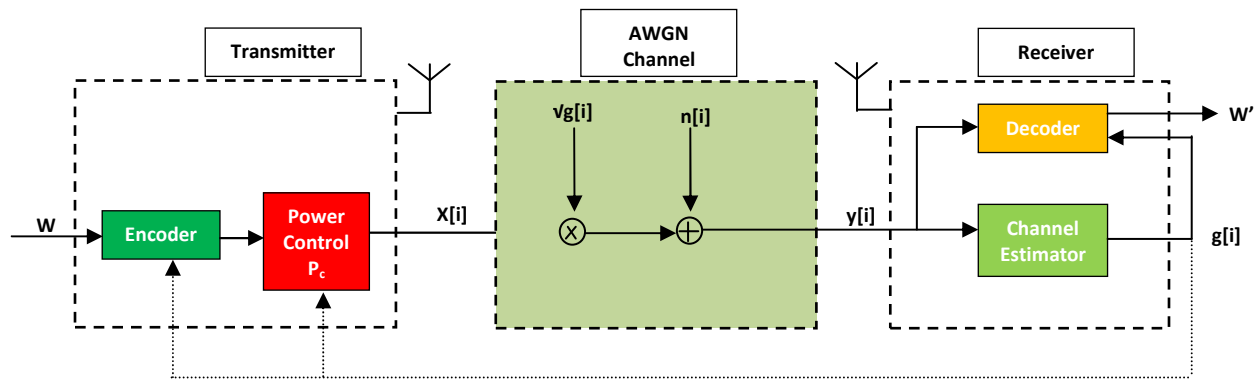
Figure 38. Model of a system with Channel Side Information (CSI) known to both transmitter and receiver: Adapted from (Goldsmith and Pravin, 1997: 1987).

5.1.   CHANNEL CAPACITY OF SPECTRUM SHARING COGNITIVE RADIO

In performance analysis of wireless communication systems, certain channel capacity indicators such as outage capacity, ergodic capacity, and effective capacity are employed to analyze the achievable rate of information transmission. The increasing demand for wireless communication and the appearance of more delay sensitive wireless applications are some of the major reasons why it is essential, as part of radio system design, to analyze the capacity limits of such systems. Estimating the system capacity limits provides information about the maximum data transmission rates with asymptotically small error probability while neglecting the delay or complexity of the encoding and the decoding units (Goldsmith, 2005:99). In this text, the ergodic capacity and the effective capacity of cognitive radio in the presence of Nakagami-*m* fading are thoroughly investigated with an assumption that the CSI is known at both ends of the channel.

That is, the achievable information rate of spectrum sharing cognitive radio systems based on ergodic and effective capacity analysis techniques.

5.1.1. Ergodic capacity of Cognitive Radio with MRC Diversity

Ergodic capacity of a wireless communication system is being investigated when the maximum achievable transmission rate is studied with an assumption that the communication duration is longer enough for the propagation to experience all channel states. As defined in (Li, 2012), it is the maximum of the long-term average achievable rate with an arbitrary small probability of error subject to constraints on the transmit power (Li, 2012: 850). According to (Goldsmith, 2005), the Shannon capacity of a fading channel with receiver CSI for an average power constraints is given in equation 5.2, which is derived by averaging the Shannon capacity for an AWGN channel of equation 5.1 over the distribution of SNR, γ given in 4.26- that is why Shannon capacity is often refer to as ergodic capacity (Goldsmith, 2005: 103 – 104; Goldsmith & Pravin, 1997: 1987; Yao & Asrar, 1993: 134).

$$C = B\log_2(1 + \gamma) \tag{5.1}$$

$$C = \int_0^\infty B\log_2(1 + \gamma)P(\gamma)\,d\gamma \tag{5.2}$$

The asymptotical analysis of Ergodic capacity of a fast-fading channel environment with long transmitted codeword has been studied in (Li, 2012), where the long-term average achievable transmission rate with arbitrary small probability of error was investigated with consideration of the transmit power constraints. That is, the author categorically investigated and analyzed the asymptotic performance of the ergodic capacity of CR system when the number of CR receive antennas is large and fading is Rayleigh (Li, 2012: 850 – 851). Also, Duan et al (2010) studied the asymmetric fading case, where the paths between SU_{tx} → SU_{rx} and SU_{tx} → PU_{rx} are Rayleigh and Nakagami fading respectively. On the other hand, this study investigates the ergodic capacity of a channel environment with SU_{tx} → SU_{rx} and SU_{tx} → PU_{rx} paths having fading components that are statistically Nakagami distribution. Therefore, with MRC in CR

systems and assuming Nakagami fading, the ergodic capacity is mathematically obtained by averaging the channel capacity over the PDFs of the channel power gains, g and h as:

$$C_{erg} = B\left[\int_{\gamma_{MRC}} \log(1 + x)P_g(g)\,P_{h_{MRC}}(h)\,dg\,dh\right] \tag{5.3}$$

Equation 5.3 is not written in closed form expression, but Li (2012) provides proof to show that the ergodic capacity of the system scales as $\ln K$ in the P_C dominated region where $P_C \ll Q$. In this text, it is also assumed that the CR secondary receivers (SU_{rx}) are equipped with large number of K receive antennas, which makes the ergodic capacity of the system to scale as a logarithmic function of K. The proof given in (Li, 2012) suggests that the variable of $\gamma_{MRC} = \sum_{i=1}^{K}\gamma_i$ for large value of K antennas can be approximated by the statistics of γ_i. Considering certain statistical properties (see *Theorem 1* in Li, 2012), the ergodic capacity with MRC as K approaches infinity scales as $\log K$ (Li, 2012: 850 – 851).

Applying ***Jensen's inequality***, equation 5.3 is less than equation 5.1 and with the interference power constraint equation 5.3 does not provide maximum capacity (Goldsmith & Pravin, 1997: 1988). In order to consider power constraint as a factor of achieving maximum information rate, it is assumed that the average interference power constraint Q is the average received power constraint at the PUrx and that channels ***g*** and ***h*** are independent to each other, the capacity per unit Hz as a function of Q becomes (Duan et al., 2010:2):

$$\frac{C_{erg}}{B} = E\left[\log\left(1 + \frac{hP(h,g)}{N_oB}\right)\right] \tag{5.4}$$

Considering the average interference power constraint, Q at the primary receiver, the ergodic capacity in 5.4 above can be formulated as an optimization problem:

$$\underset{p(h,g)}{maximize} \quad E\left[\log\left(1 + \frac{hP(h,g)}{N_oB}\right)\right] \tag{5.5a}$$

$$subject\ to \quad E_{h,g}[P(h,g)g] \leq Q \tag{5.5b}$$

With the help of *theorem 7.16* in Foulds (1981:285 – 286), the maximized ergodic capacity subject to power constraint can be obtained. Therefore, applying Lagrangian optimization technique, the solution to the above optimization problem can be simplified as:

$$P(h, g) = \left(\frac{1}{\lambda g} - \frac{N_o B}{h} \right)^+ \qquad (5.6)$$

Where $(.)^+$ indicates max $(., 0)$ and $\lambda \geq 0$ is the Lagrangian multiplier satisfying the condition $E_{h,g}[P(h, g)g] - Q = 0$

Therefore, putting the PDFs of the channel gains in 4.25 and 4.27 (with MRC) into 5.5a with the interference power constraint, transmission experiencing Nakagami fading on both paths, and the SUrx is implemented with MRC; the capacity per unit bandwidth (Hz) of the cognitive radio link can be expressed as follows (*see proof in* Appendix I), where $\gamma_o = 1/\lambda$ for normalized case:

$$\frac{C_{erg}}{B} = \frac{\frac{m_o{}^{m_o} m^{Km}(Km - 1)!}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}}}{\left(\frac{m}{\gamma_o}\right)^{Km}} \sum_{\alpha=0}^{Km-1} \frac{1}{\alpha!} \frac{\left(\frac{m}{\gamma_o}\right)^\alpha \Gamma(m_o + \alpha)}{m_o \left(\frac{m}{\gamma_o} + m_o\right)^{m_o+\alpha}} \,_2F_1\left(1, m_o + \alpha; m_o + 1; \frac{m_o}{\frac{m}{\gamma_o} + m_o}\right) \quad (5.7)$$

Since the above ergodic capacity expression is subject to the average interference power constraint as shown in the optimization problem described by (5.5a) and (5.5b), the average interference power constraint can be simplified by inserting the PDFs of the channel power gains in equations 4.25 and 4.27 into the power constraint expression in (5.5b), therefore, the obtainable average interference power constraint, Q for normalized case is (*see proof in* Appendix II):

$$\frac{Q}{N_o B} = \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \frac{\gamma_o{}^{m_o+1}\Gamma(Km + m_o)}{(m_o\gamma_o + m)^{Km+m_o}} \left[\frac{\,_2F_1\left(1, Km + m_o; m_o + 1; \frac{m_o\gamma_o}{m_o\gamma_o + m}\right)}{m_o{}^3} \right.$$

$$\left. - \frac{1}{(m_o + 1)} \,_2F_1\left(1, Km + m_o; m_o + 2; \frac{m_o\gamma_o}{m_o\gamma_o + m}\right) \right] \qquad (5.8)$$

5.1.2. EFFECTIVE CAPACITY OF COGNITIVE RADIO WITH MRC DIVERSITY

The effective capacity of a channel is also an important link model in radio channel analysis, just as ergodic capacity. The idea of effective capacity (E_C) for time varying fading channels was first proposed in (Wu & Rohit, 2003) as a wireless link model that characterizes radio channel in

terms of quality of service (QoS) metrics (such as delay, data rate, and delay-violation probability) as opposed to link characterization based on fluctuations in the amplitude of radio signals. In other words, ergodic capacity measures achievable information rate at the *physical layer* while effective capacity has been proposed as a *link-layer* channel capacity model (Wu & Rohit, 2003:630), which provides support for QoS requirements for analyzing the effect of queuing on channel capacity. Therefore, in this study, the reliability on the capacity information obtained from effective capacity model is compared to that of ergodic capacity.

The PDFs of the channel power gains *g* and *h* are given in equations 4.25 and 4.26 respectively and with MRC at the SU_{rx}, the PDF of the channel gain *h* is given by 4.27. The assumption that the CSI is available at both ends of the link and defining the power allocation policy to satisfy the transmit power constraint Q, the achievable effective capacity can be obtained. Also, it is assumed that the transmission mode of the SU_{tx} satisfies a statistical delay QoS constraint exponent, θ expressed as (Wu & Negi, 2003):

$$\theta = -\lim_{\varphi \to \infty} \frac{\ln\left(\Pr\{q(\infty) > \varphi\}\right)}{\varphi} \tag{5.9}$$

Where q(*x*) is the probability that the queue length of the transmit buffer exceeds a certain threshold, φ, and decays exponentially as a function of *x*. Given this QoS constraint, the maximum arrival-rate of the SU can be obtained. In the assumed CR system, there is no delay constraint when θ → 0 but a strict delay constraint exists when θ → ∞ (Musavian, Leila et al., 2010: 1700; Musavian & Sonia, 2010: 1055).

Considering the probability of the service queue length defined in the delay QoS constraint of equation 5.9, the E_C can be obtained by assuming that the maximum achievable service rate $R[i]$ is a discrete time service rate, which is stationary and ergodic. Assuming the channel in **Figures 37** and **38** experiences block fading, which is also ergodic and stationary, the service rate of block *i* can be expressed as (Musavian, Leila et al., 2010: 1700; Musavian & Sonia, 2010: 1055):

$$\{R[i], i = 1, 2, 3, \ldots \ldots\}$$

Then the effective capacity without block fading can be expressed as:

$$E_c(\theta) = -\lim_{t \to \infty} \frac{1}{\theta t} \ln\left(\varepsilon\{e^{-\theta \sum_{i=1}^{n} R[i]}\}\right) \tag{5.10}$$

For an uncorrelated sequence $i = 1, 2... k$, the effective capacity E_C for a block fading channel becomes:

$$E_c(\theta) = -\frac{1}{\theta} \ln\left(\varepsilon\{e^{-\theta R[i]}\}\right) \tag{5.11}$$

Since transmit power constraint with respect to optimum power allocation is very vital in CR systems, the maximum achievable effective capacity is subject to the power constraints conditions and the QoS constraint exponent θ expressed in equation 5.9. Therefore, the maximum achievable instantaneous service rate R[i] of block i becomes:

$$R[i] = TB \ln\left(1 + \frac{p(\theta,h,g)g}{NoB}\right) \tag{5.12}$$

Where g denotes the channel power gain from the SU_{tx} to the PU_{rx}, h is the channel power gain from the SU_{tx} to the SU_{rx}, T is the time duration of a block, and B represents the channel bandwidth. Following this expression that the maximum effective capacity is subject to power and QoS constraints, the maximum effective capacity can be expressed as an optimization problem. Using Lagrangian optimization technique, the optimization problem of the effective capacity with the constraints can be formulated as follows (Musavian & Sonia, 2010: 1056):

$$E_C = maximize\left\{-\frac{1}{\theta} \ln\left(\varepsilon_{h,g}\left\{e^{-TB\ln\left(1+\frac{P(\theta,h,g)g}{NoB}\right)}\right\}\right)\right\} \tag{5.13}$$

$s.t \;\; \varepsilon_{h,g}\{P(\theta,h,g)g\} \le Q$

$P(\theta,h,g)g \ge 0$

Since function ln(x) is a monotonically increasing function of x, the solution to the optimization problem in 5.13 is the same as the solution for the minimization problem expressed as:

$$minimize\left\{\varepsilon_{h,g}\left\{\left(1 + \frac{P(\theta,h,g)g}{NoB}\right)^{-\theta TB}\right\}\right\} \tag{5.14}$$

$s.t \;\; \varepsilon_{h,g}\{P(\theta,h,g)g\} \le Q$

$P(\theta,h,g)g \ge 0$

The solution to the optimization problem of equation 5.14 can be obtained using the Lagrangian optimization technique:

$$\mathcal{L}(P(\theta,h,g),\lambda_0) = \lambda_0\big(\varepsilon_{h,g}\{P(\theta,h,g)g\} - Q\big) + \varepsilon_{h,g}\left\{\left(1+ \frac{P(\theta,h,g)g}{NoB}\right)^{-\theta TB}\right\} \tag{5.15}$$

Considering the fact that the optimum allocation should satisfy $\frac{\partial \mathcal{L}(P(\theta,h,g),\lambda_o)}{\partial P(\theta,h,g)} = 0$ then

$$P(\theta,h,g) = NoB\left[\frac{\beta^{\frac{1}{1+\alpha}}}{g^{\frac{1}{1+\alpha}}h^{\frac{\alpha}{1+\alpha}}} - \frac{1}{h}\right]^{+} \text{where } \alpha = \theta TB, \beta = \frac{\gamma_o\alpha}{NoB}, [x]^{+} \text{ denotes } \max\{0, x\}, \gamma_o = \frac{1}{\lambda_o},$$

and $\lambda_o \geq 0$ is the Lagrange multiplier associated to the average interference power constraint, and γ_o must be determined in order to satisfy the interference power constraint with equality. Therefore, the optimum power allocation policy can be formulated as thus:

$$P(\theta,h,g) = \begin{cases} NoB\left(\dfrac{\beta^{\frac{1}{1+\alpha}}}{g^{\frac{1}{1+\alpha}}h^{\frac{\alpha}{1+\alpha}}} - \dfrac{1}{h}\right) & if \ g \leq \beta h \\ \\ 0 & otherwise \end{cases} \tag{5.16}$$

The optimum cutoff ratio γ_o can be obtained by integrating the optimum power allocation function over the PDFs of the channel power gains, g and h given in (4.25) and (4.27):

$$\frac{Q}{NoB} = \int_0^\infty \int_0^{\beta h} \left(\beta^{\frac{1}{1+\alpha}}\left(\frac{g}{h}\right)^{\frac{\alpha}{1+\alpha}} - \frac{g}{h}\right) * P_g(g)P_{hMRC}(h)\, dg\, dh \tag{5.17}$$

Let random variable v be defined as $v = \frac{g}{h}$ and using the fact that the distribution of the ratio between two Gamma distributed random variables with parameters Km and m_o is a beta prime distribution with parameters Km and m_o (Weisstein, 2012), the distribution of the random variable v becomes equation 5.18 where $z = \frac{Km}{m_o}$ and B(.,.) is a beta function:

$$f_v(v) == \frac{z^{Km}}{B(Km,m_o)}\frac{v^{m_o-1}}{(Z+v)^{m_o+Km}} \tag{5.18}$$

Therefore, with this joint distribution defined above (*see proof* in APPENDIX III), the integrals in equation 5.17 above can be evaluated to determine the cutoff ratio γ_o as follows:

$$\frac{Q}{NoB} = \frac{z^{Km-1}}{B(Km, m_o)} \int_0^\beta \left(\beta^{\frac{1}{1+\alpha}} v^{\frac{\alpha}{1+\alpha}} - v \right) \frac{v^{m_o-1}}{(v+Z)^{Km+m_o}} \, dv \tag{5.19}$$

Then with the help of the Gauss's Hypergeometric function defined in (Musavian & Sonia, 2010: 1056 eq. 5; Alouini & Andrea, 1999:1179) to express the integrals in 5.17, the expression for the average interference power is obtained as (*see* APPENDIX IV *for proof*):

$$\frac{Q}{NoB} = \frac{\beta^{m_o+1}}{B(Km, \, m_o)Z^{m_o} \left(m_o + \frac{\alpha}{1+\alpha} \right)} \, {}_2F_1 \left(Km + m_o; \, m_o + \frac{\alpha}{1+\alpha} \, ; \, m_o + \frac{\alpha}{1+\alpha} + 1, -\frac{\beta}{Z} \right)$$

$$- \frac{\beta^{m_o+1}}{B(Km, \, m_o)(m_o+1)Z^{m_o}} \, {}_2F_1 \left(Km + m_o; \, m_o + 1 \, ; \, m_o + 2, -\frac{\beta}{Z} \right) \tag{5.20}$$

In order to obtain the closed form expression, the Gauss Hypergeometric function, ${}_2F_1(.)$ can be transformed using equation 9.131 in (Gradshteyn and Ryzhik, 2007: 1008) given as follows:

$$F(\alpha, \beta; \, \gamma; z) = (1-z)^{-\alpha} F \left(\alpha, \gamma - \beta; \, \gamma; \frac{z}{z-1} \right)$$

Applying this Gauss Hypergeometric *transformation* formula, the closed form expression the power constraint in equation 5.20 above was obtained as:

$$\frac{Q}{NoB} = \frac{\beta^{m_o+1} Z^{Km}}{B(Km, \, m_o) \left(m_o + \frac{\alpha}{1+\alpha} \right) (z+\beta)^{Km+m_o}} \, {}_2F_1 \left(Km + m_o; \, 1 \, ; \, m_o + \frac{\alpha}{1+\alpha} + 1, \frac{\beta}{\beta+z} \right)$$

$$- \frac{\beta^{m_o+1} Z^{Km}}{B(Km, \, m_o)(m_o+1)(z+\beta)^{Km+m_o}} \, {}_2F_1 \left(Km + m_o; \, m_o + 1 \, ; \, m_o \right.$$

$$\left. + 2, \frac{\beta}{\beta+z} \right) \tag{5.21}$$

The effective capacity of a spectrum sharing cognitive radio channel as a function of the delay exponent can be formulated in terms of power allocation constraint as in 5.22 below, where $p := p(\theta, h, g)$:

$$E_c(\theta) = -\frac{1}{\theta} \ln \left(\varepsilon_v \left\{ \left(1 + \left[\beta^{\frac{1}{1+\alpha}} v^{-\frac{1}{1+\alpha}} - 1 \right]^+ \right)^{-\alpha} \right\} \right)$$

$$= -\frac{1}{\theta} \ln \left(\frac{z^{Km-1} \beta^{-\frac{\alpha}{1+\alpha}}}{B(Km, m_o)} \int_0^\beta \frac{v^{m_o - \frac{1}{1+\alpha}}}{(v+Z)^{Km+m_o}} \, dv \right.$$

$$\left. + \frac{z^{Km-1}}{B(Km, m_o)} \int_0^\beta \frac{v^{m_o-1}}{(v+Z)^{Km+m_o}} \, dv \right) \tag{5.22}$$

Using the fact that $\int_0^\infty f_v(v)dv = 1$, the closed form expression for the effective capacity of the secondary user operating in a spectrum sharing (*underlay*) cognitive radio channel subject to interference power constraint has been expressed as equation 5.23 below (*see* APPENDIX V *for proof*), where m_o is the fading parameter of the path between SU and PU, Km is the fading parameter with MRC K antennas, and $_2F_1(.)$ is a Gaussian hypergeometric function whose general integral representation is given in (*equation* 9.111 in Gradshteyn and Ryzhik, 1994: 1066).

$$E_c(\theta) = - \frac{1}{\theta} \ln \left(\frac{\beta^{m_o}}{B(Km, \, m_o)Z^{m_o}\left(m_o + \frac{\alpha}{1+\alpha}\right)} \, _2F_1\left(Km + m_o; \, m_o + \frac{\alpha}{1+\alpha} \, ; \, m_o + \frac{\alpha}{1+\alpha} \right. \right.$$

$$\left. + 1, -\frac{\beta}{Z}\right) + 1$$

$$\left. - \left[\frac{\beta^{m_o}}{B(Km, \, m_o)m_o Z^{m_o}} \, _2F_1\left(Km + m_o; \, m_o \, ; \, m_o + 1, -\frac{\beta}{Z}\right)\right] \right) \qquad (5.23)$$

The closed form expression for the effective capacity in equation 5.23 above can be obtained by using the same Gauss Hypergeometric *transformation* formula (*equation 9.131* in Gradshteyn and Ryzhik, 2007: 1008) on the previous page. The resultant effective capacity after this transformation was obtained as given in below:

$$E_c(\theta) = - \frac{1}{\theta} \ln \left(\frac{\beta^{m_o} \, Z^{Km}}{B(Km, \, m_o)(z+\beta)^{Km+m_o}\left(m_o + \frac{\alpha}{1+\alpha}\right)} \, _2F_1\left(Km + m_o; \, 1; \, m_o + \frac{\alpha}{1+\alpha} \right. \right.$$

$$\left. + 1, \frac{\beta}{Z+\beta}\right) + 1$$

$$\left. - \left[\frac{\beta^{m_o} \, Z^{Km}}{B(Km, \, m_o)m_o(z+\beta)^{Km+m_o}} \, _2F_1\left(Km + m_o; \, 1; \, m_o + 1, \frac{\beta}{Z+\beta}\right)\right] \right) \qquad (5.24)$$

5.2.   NUMERICAL RESULTS, ANALYSIS AND DISCUSSION

In the preceding section, the ergodic capacity and the effective capacity of Nakagami-*m* CR fading channels subject to average interference power constraint have been mathematically illustrated, and in addition, the effective capacity is subject to delay QoS constraint and assumed Rayleigh block fading channel. This section presents the numerical results of the maximum achievable information transmission rates. In this analysis, it is assumed that the *underlay* spectrum-sharing approach is being used, where the secondary users share the spectrum with the primary user such that the transmission of the SU does not interfere with that of the PU. The results for the ergodic capacity and effective capacity analyses are normalized by assuming that $N_oB = 1$ with various values of average interference power, Q (dB), and different values of delay QoS constraints in the case of effective capacity.
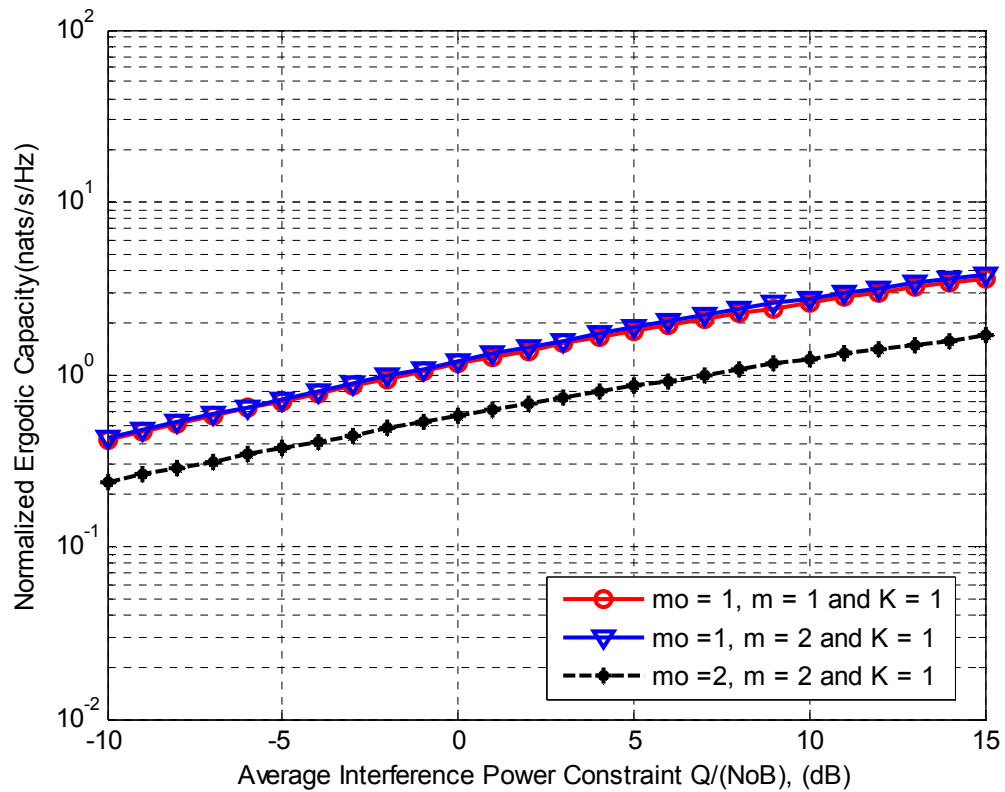

Figure 39. Normalized ergodic capacity per unit bandwidth for the same degree of fading on both $SU_{tx} \rightarrow SU_{rx}$ and $SU_{tx} \rightarrow PU_{rx}$ paths $m_o = m = 1$ and $m_o = m = 2$ subject to transmit power constraint, Q, without **MRC** (K = 1) at the SU_{rx}.

The ergodic capacity of the SU_{rx} with K branch increases when the degree of Nakagami-m fading parameter m increases. The results in this analysis also reveal that interference to the primary system increases when the Nakagami fading along SU_{tx} → PU_{rx} path is less severe (m_o decreases). But when the two paths, SU_{tx} → PU_{rx} and SU_{tx} → SU_{rx} experience equal degree of Nakagami-m fading ($m_o = m$), the CR channel capacity increases proportionally to the number of MRC diversity K antennas at the secondary user-receiver. The benefit of MRC in CR system design can be inferred from comparing **Figure 39** and **40** below. Without MRC (K = 1) and the average interference power Q = -5dB in **Figure 39**, the maximum achievable capacity is 0.4 nats/s/Hz while with MRC (K = 2) in **Figure 40**, the capacity under -5dB power constraint gives a capacity of 1 nats/s/Hz. Also, in the same **Figure 40**, using 0dB power constraint as reference point, it is obvious that increasing the number of MRC antennas (K = 2 → 4 → 8) at the SUrx, yields an increase in ergodic capacity; although, not so significant or much increase, but there is a gain in capacity when the number antennas is increased.
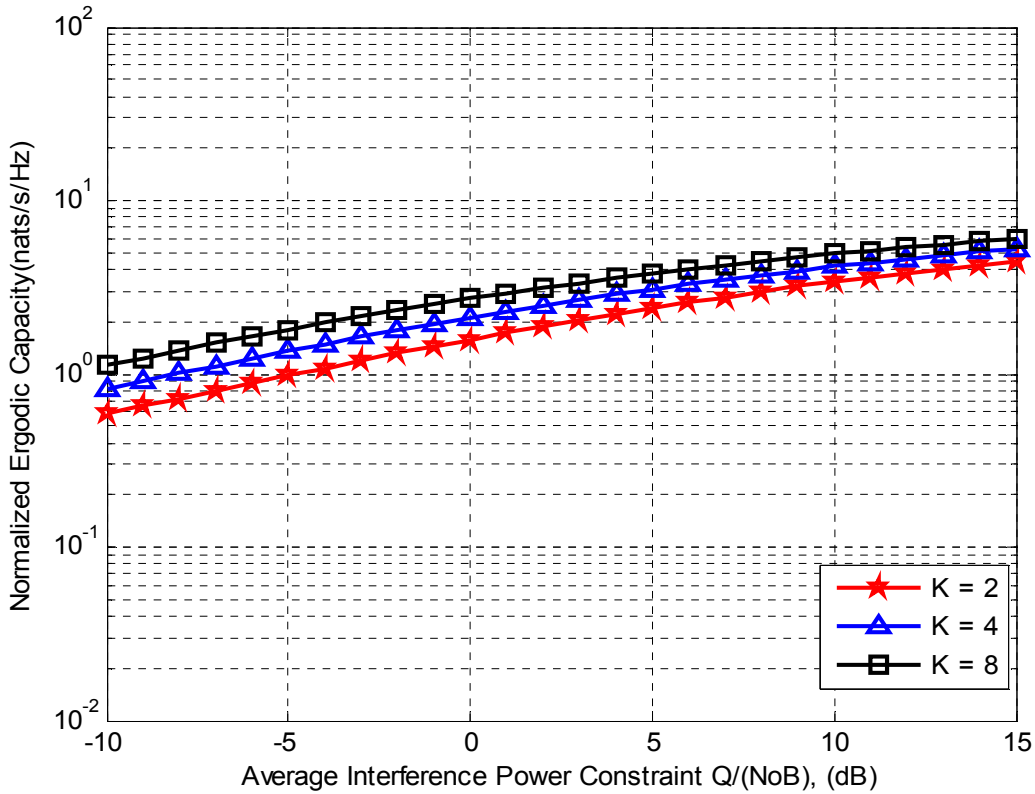

Figure 40. Normalized ergodic capacity per unit bandwidth for same degree of fading on both SU_{tx} → SU_{rx} paths m_o = m = 1 and SU_{tx} → PU_{rx} subject to transmit power constraint, Q (dB), and with **MRC** (K = 2, 4, and 8) at the SU_{rx}.

In analyzing the CR effective capacity, it is further assumed that NoB = 1 and TB =1, thereby normalizing the effective capacity results. **Figures 41a** and **41b** illustrate the normalized achievable effective capacity of CR channel under different average interference power, Q with assumed delay QoS exponent θ = *0.01* (1/nat). From these Figures, it is apparent that the effective capacity increases with increase in number of MRC antennas K at the SU$_{rx}$. In Figure 41a, the increase in capacity is significant when the Nakagami-*m* parameter of fading *m* = 0.5 along the SU$_{tx}$ → SU$_{rx}$ path. On the contrary, in **Figure 41b**, with better secondary channel, *m* is Rayleigh and increase in MRC antennas K at the secondary receiver, the increase in effective capacity is less significant but more than when *m* = 0.5. Taking 0dB interference power as reference point in **Figure 41a**, when m_o = *m* = 0.5 and K = 2, the achievable transmission rate under 0.01 delay QoS exponent is 1.7 nats/s/Hz, and at the same interference power reference in **Figure 41b** where K = 2 and m = Rayleigh, the achievable capacity is 1.8 nats/s/Hz due to better channel.

Figure 41a. Normalized effective capacity per unit bandwidth subject to transmit power constraint, Q (dB), with **MRC** (K = 1, 2, and 8) at the SU$_{rx}$ and delay QoS exponent, θ = 0.01 when SU$_{tx}$ → SU$_{rx}$ Nakagami path *m* = 0.5 while SU$_{tx}$ → PU$_{rx}$ Nakagami path m_o = 0.5.

Figure 41b. Normalized effective capacity per unit bandwidth subject to transmit power constraint, Q (dB), with **MRC** (K = 1, 2, and 8) at the SU$_{rx}$ and delay QoS exponent, θ = 0.01 when SU$_{tx}$ → SU$_{rx}$ Nakagami path *m* = *Rayleigh* while SU$_{tx}$ → PU$_{rx}$ Nakagami path m_o = 0.5.

So far, it has been buttressed with **Figure 41a** and **41b** that implementing MRC antennas diversity at the secondary receiver in CR systems can increase the maximum achievable information transmission rate under Rayleigh block fading CR channels. **Figure 42** illustrates the normalized effective capacity of CR channel when the Nakagami fading parameters m_o (SU$_{tx}$ $\rightarrow$ PU$_{rx}$) is Rayleigh and that of the secondary channel $m = 0.5$. Therefore, when the primary or interference channel gets better, the capacity of CR decreases significantly compare to the achievable capacity in **Figure 41b** where the secondary channel is better. Although, when the interference channel is better than the CR channel, the capacity can still be increased by increasing the MRC antennas but not as significant as the achievable capacity when the secondary channel is better than the primary channel. **Figure 41b** and **Figure 43** show improved CR capacity when CR channel improves, m is Rayleigh and $m = 3$ respectively.
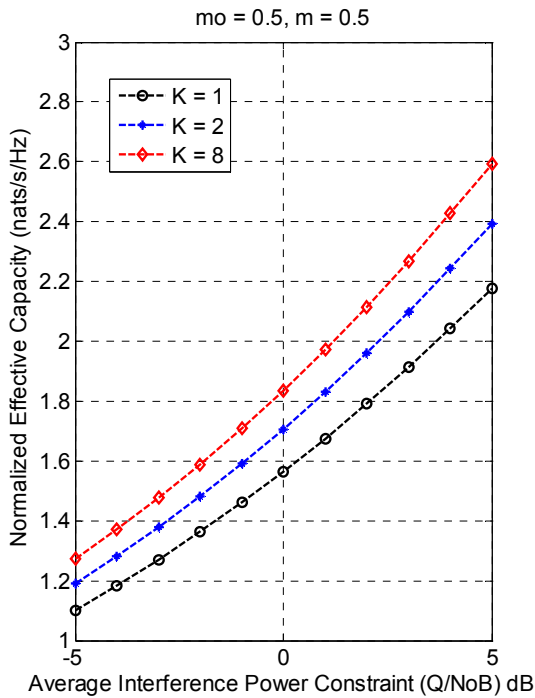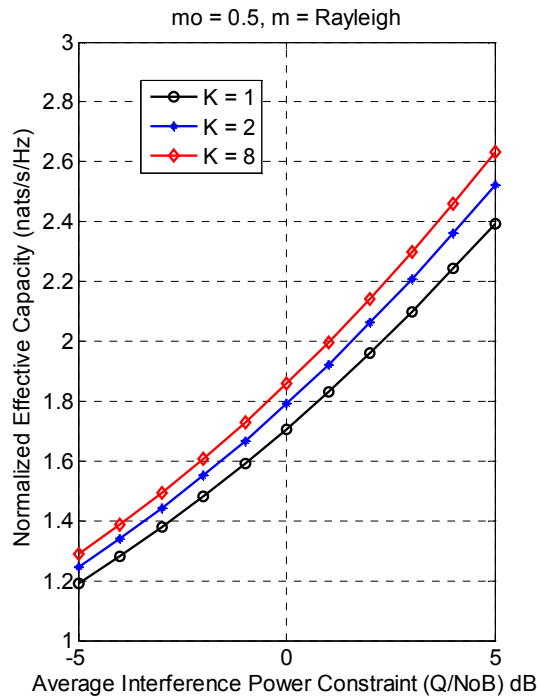

Figure 42. Normalized effective capacity per unit bandwidth subject to transmit power constraint, Q (dB), with **MRC** (K = 1, 2, and 8) at the SU$_{rx}$ and delay QoS exponent, $\theta = 0.01$ when SU$_{tx}$ $\rightarrow$ PU$_{rx}$ Nakagami path has fading parameter m_o is Rayleigh while SU$_{tx}$ $\rightarrow$ SU$_{rx}$ has Nakagami path fading parameter $m = 0.5$.
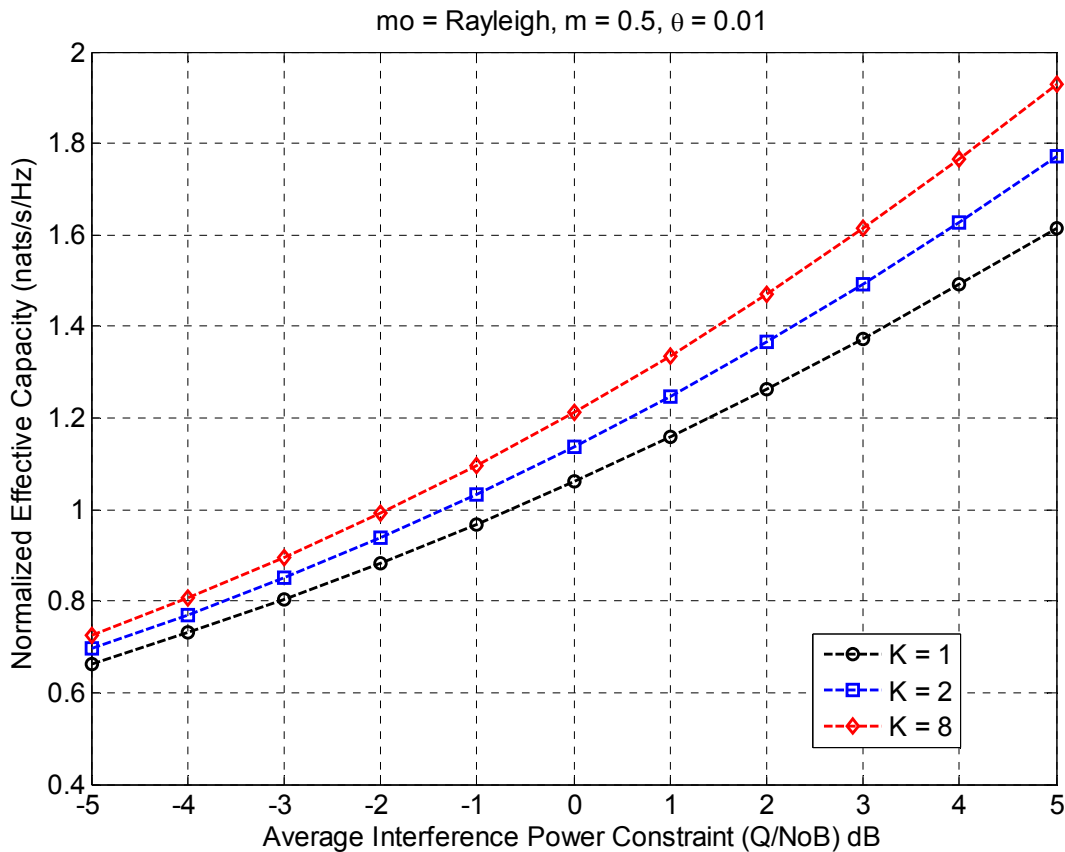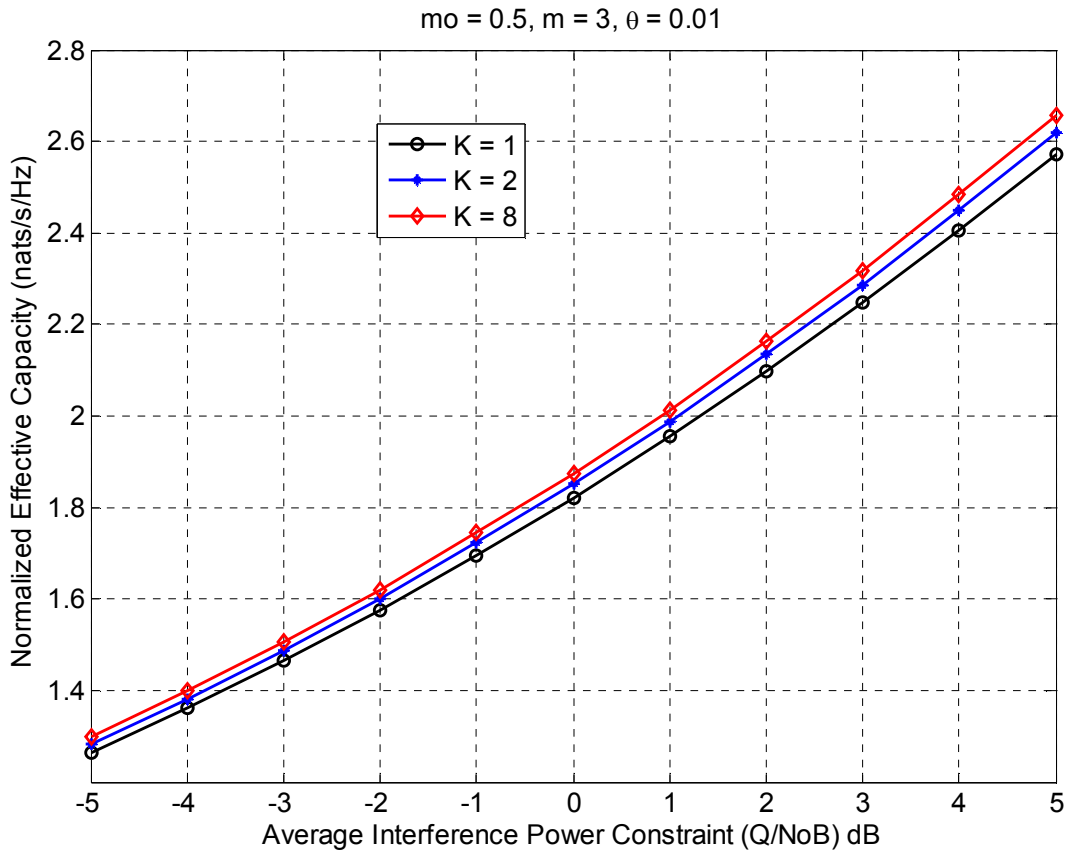
mo = 0.5, m = 3, θ = 0.01

Figure 43. Normalized effective capacity per unit bandwidth subject to transmit power constraint, Q (dB), with **MRC** (K = 1, 2, 8) at the SU_{rx} and delay QoS exponent, $\theta = 0.01$ when $SU_{tx} \rightarrow SU_{rx}$ Nakagami path parameter is constant at $m = 3$ while $SU_{tx} \rightarrow PU_{rx}$ Nakagami path parameter increases as $m_o = 0.5$, Rayleigh, 2, and 4.

At this point, it is worth to point out that in **Figure 43**, even though the CR channel gets better, the increase in capacity is not very significant due to transmit power limitation that hinders the use of all the opportunity to transmit. It is also ideal to analyze the effective capacity of CR system under different delay QoS exponent. The numerical results of $E_c(\theta)$ against various delay QoS exponents illustrated in **Figure 44** at Q = - 5dB. It is obvious that the effective capacity of spectrum sharing CR decreases as the delay QoS exponent θ increases. Hence, increasing the delay QoS exponent reduces the maximum information transmission rate of CR systems. The evidence from this result justifies the capability of CR systems in satisfying various QoS specifications for different CR users or different wireless applications such as real -time or delay-sensitive applications (voice and video transmission etc). In other words, different QoS requirements for different users or applications can be supported by cognitive radio under the same power constraint for all CR users.
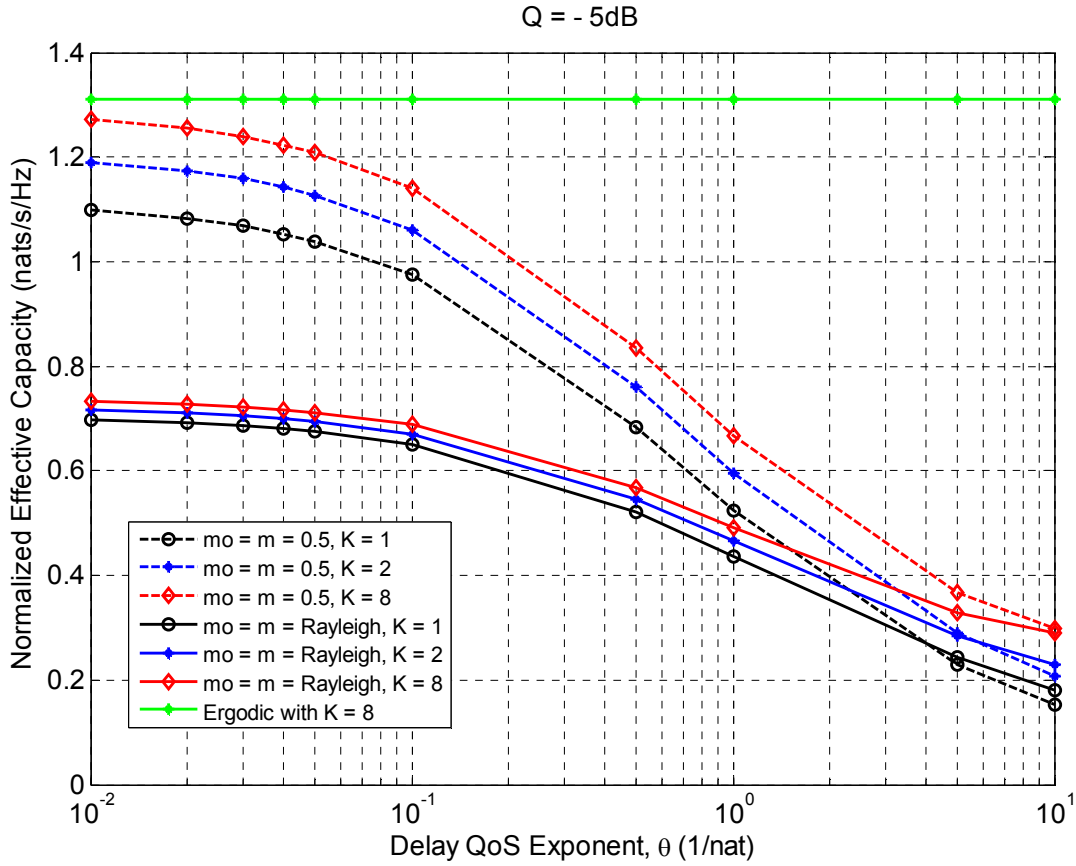
Figure 44: Normalized effective capacity versus ergodic per unit bandwidth subject to transmit power constraint, Q (dB), with **MRC** (K = 1, 2, and 8) at the SU_{rx} and various delay QoS exponent values at Q = -5dB.

Lastly, in comparison with the ergodic capacity at the same Q = -5dB, it is apparent that as the delay QoS exponent decreases, the effective capacity approaches ergodic capacity. Although, delay QoS exponent is not defined for ergodic capacity since the model does not take delay QoS constraint into consideration, but we can comparatively analyze ergodic and effecitve capacity on the same logarithm scale. With MRC, the increase in effective capacity is more significant than in ergodic capacity, but when the delay QoS exponent decreases, for instance to $\theta = 0.01$ at Q = -5dB and K = 8 in **Figure 44**, the effective capacity decreases and approaches the same value as ergodic capacity. And when the QoS exponent increases towards infinity, the effective capacity approaches outage capacity.

6.   CONCLUSION AND FUTURE WORK

Cognitive radio access techniques promise efficient use of the scarce spectrum resource and peaceful coexistence of dissimilar radio technologies on the radio portion of the electromagnetic spectrum. In order to achieve these benefits, studies on spectrum sharing technique is very paramount. In contribution to the realization of an operational cognitive radio technology and to the body of knowledge in the field of communication engineering, this thesis was extensively and independently carried out to derive the mathematical expressions for analyzing the effective and ergodic capacity of spectrum sharing cognitive radio channel. It was assumed that the spectrum sharing technique is the *underlay* scheme and that multiple antennas are implemented at the secondary receiver based on MRC diversity technique and the channel fading is characterized as Nakagami-*m* fading. As an independent work, the ergodic and effective capacity mathematical expressions were derived to investigate the maximum achievable information transmission rates subject to interference power constraints.

The mathematical approximations for the ergodic capacity and the effective capacity have been obtained, and the maximum achievable capacity of spectrum sharing cognitive radio has been numerically illustrated. The ergodic capacity model has been used to investigate the CR capacity at the PHY layer while the effective capacity model fostered CR system capacity investigation at the link layer. In both cases, the channel from the secondary transmitter to the secondary receiver ($SU_{tx} \rightarrow SU_{rx}$) and from the secondary transmitter to the primary receiver ($SU_{tx} \rightarrow PU_{rx}$) experience Nakagami fading, and the secondary receiver has MRC antennas implemented. The numerical results show that with MRC at the SUrx, there is an increase in both effective and ergodic capacity of the CR system, and that MRC could be used to combat the unwanted effects of fading in CR systems.

In comparison, effective capacity model is suitable for capacity analysis when the correlation between the maximum achievable information transmission rate and the specified QoS conditions is to be established. Effective capacity tells us the link capacity under various QoS specifications, and the results obtained under effective capacity analysis support the notion that CR networks have the capability to support different QoS requirements for different users and

wireless applications. But on the other hand, ergodic capacity analysis does not take QoS constraint into consideration. Since cognitive radio must be implemented with interference limit consideration, both the effective capacity and ergodic capacity analyses are subject to average interference power constraint. Hence, while the effective capacity is subject to both interference power and satisfying the delay QoS exponent, the ergodic capacity is subject to only the interference power limit. In both capacity analyses, it has been demonstrated that the secondary user can share the same spectrum with the primary user provided the average interference power constraint is defined such that harmful interference to the PU can be avoided. The effective and ergodic capacity of the SU depends heavily on the magnitude of the interference power limit; each capacity increases with increase in average interference power. However, the interference power can be kept low to avoid unwanted interference, and by increasing the number of MRC antennas at the secondary user, both effective and ergodic capacity can be increased even without increasing the average interference power.

The possible future work is to investigate the BER of CR with MRC over Nakagami-m fading environment. Similar comparative study can be done on the performance analysis of ergodic and effective capacity of CR using different modulation schemes such as BPSK, QAM-64 et cetera. Future works can also include the investigation of the performance of CR systems when transmit power and interference power constraints are considered. Investigating the average power expenditure to achieve the effective capacity of cognitive radio with MRC is another research gap. Also, to practically verify the numerical results obtained in this thesis, a cognitive radio test bed can be implemented in future research. How about investigating a hybrid capacity analysis, whereby the effective and ergodic capacity are mathematically combined to form one model to achieve a cross-layer throughput or capacity model for cognitive radio? Cognitive radio has various research gaps because it is the future of radio communication.

7. REFERENCES

Akyildiz F. Ian et al. (2006). *Next Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey*. Elsevier B. V: Computer Networks 50, 2006, 2127 – 2159

Alouini, Mohamed-Slim and Andrea Goldsmith (1997). *Capacity of Nakagami Multipath Fading Channels*. Vehicular Technology Conference, 1997, IEEE 47th, Vol. 1, pp. 358 – 362

Alouini, Mohamed-Slim and Andrea Goldsmith (1999). *Capacity of Rayleigh Fading Channels Under Different Adaptive Transmission and Diversity-Combining Techniques*. IEEE Transactions on Vehicular Technology, Vol. 48, No.4, pp. 1165 – 1181

Anderson, J. B., et al. (1995). *Propagation Measurements and Models for Wireless Communications Channels*. IEEE Communications Magazine, vol. 33, No. 1, pp. 42 – 49

Bhattacharya, C. et al. (2009). *Generation of Nakagami-m Statistics by Monte Carlo Method for Fading Communication Channels*. India Conference (INDICON), 2009 Annual IEEE Conference Publications, pp. 1 – 4

Brennan, D. G. (2003). *Linear Diversity Combining Techniques*. Proceedings of the IEEE, Vol. 91, No. 2, February, 2003.

Cabric D. et al. (2004). *Implementation Issues in Spectrum Sensing for Cognitive Radios*. Proceedings of 38th Asilomar Conference on Signals, Systems and Computers, 2004, pp. 772 – 776

Duan Ruifeng et al. (2010). *Capacity for Spectrum Sharing Cognitive Radios with MRC Diversity at the Secondary Receiver under Asymmetric Fading*. IEEE Communications Society, Global Telecommunications Conference (GLOBECOM 2010), pp. 1 -5

Foulds L. R. (1981). *Optimization Techniques: An Introduction*. Springer – Verlag Inc, New York, U.S.A. ISBN 0-387-90586-3

FCC Spectrum Policy Task Force (2002). *Report of the Spectrum Efficiency Working Group*. Technical Report 02 – 135. [Online]. Accessed: 14/07/2012. Available: http//www.fcc.gov/sptf/files/SEWGFinalReport_1.pdf

Goldsmith Andrea et al. (2009). *Breaking Spectrum Gridlock with Cognitive Radios: An Information Theoretic Perspective*. Proceedings of the IEEE, Vol. 97, No. 5, 2005.

Goldsmith J. Andrea and Pravin P. Varaiya (1997). *Capacity of Fading Channels with Channel*

Side Information. IEEE Transactions on Information Theory, Vol. 43, No. 6, pp. 1986 – 1992

Goldsmith, Andrea (2005). *Wireless Communications.* Cambridge University Press, Cambridge, New York, U.S.A. ISBN-13 978-0-521-83716-3

Gradshteyn, I.S & Ryzhik, I.M (2007). Table of Integrals, Series, and Products. 7th Edition, Elsevier Academic Press: ISBN-13: 978-0-12-373637-6

Gradshteyn, I.S & Ryzhik, I.M (1994). Table of Integrals, Series, and Products. 5th Edition, Elsevier Academic Press: ISBN-13: 978-0-12-373637-6

Gustafsson F., et al., (2010). *Signal Processing.* Studentlitteratur: Holmbergs i Malmö AB, Sweden ISBN 978-91-44-05835-1

Jakes C. Williams, ed. (1974). *Microwave Mobile Communications.* The Institute of Electrical Electronics Engineers, Inc., Wiley-Interscience, U.S.A

Haykin Simon (2005). *Cognitive Radio: Brain-Empowered Wireless Communications.* IEEE Journal on Selected Areas in Communications, Vol. 23, No. 2, February 2005.

Jiang T., D. Grace (2011). *Efficient Exploration in Reinforcement Learning-based Cognitive Radio Spectrum Sharing.* IET Commun., 2011, Vol. 5, Iss. 10, pp. 1309 - 1317

Lundén, Jarmo (2009). *Spectrum Sensing for Cognitive Radio and Radar Systems.* Dissertation: Department of Signal Processing and Acoustics, Helsinki University of Technology, Espoo, Finland. Multiprint Oy, ISBN 978-952-248-054-5

Liang, Ying-Chang et al. (2011). *Cognitive Radio Networking and Communications: An Overview.* IEEE Transactions on Vehicular Technology, Vol. 60, No. 7, September, 2011.

Li, Dong. (2012). *Performance Analysis of MRC Diversity for Cognitive Radio Systems.* IEEE Transactions on Vehicular Technology, Vol. 61, No. 2, February, 2012.

Maliatsos et. al., (2006). *The Power Delay Profile of the Mobile Channel for Above the Sea Propagation.* Vehicular Technology Conference (VTC Fall), 2006 IEEE 64th;pp. 1 - 5

Mekkanen, Mike. (2008). *Spectrum Sensing Techniques in Cognitive Radio: Cyclostationary Method.* Master's Thesis. Telecommunication Engineering Group, Faculty of Technology, University of Vaasa, Finland.

Mitola, Joseph et al., (1999). *Cognitive Radio: Making Software Radios more personal.* IEEE Personal Communication, Vol. 6, No. 4, pp. 13 – 18.

Mitola, Joseph (2000). *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio.* Dissertation: Doctor of Technology, Royal Institute of Technology (KTH), Stockholm, Sweden.

Mitola, Joseph (1999). *Cognitive Radio: Model-Based Competence for Software Radios.* Thesis: Licentiate of Technology, Royal Institute of Technology (KTH), Stockholm, Sweden.

Mitra K., Sanjit (1999). *Digital Signal Processing: A Computer-Based Approach.* McGraw-Hill, Series in Electrical and Computer Engineering, ISBN 0-07-042953-7

Musavian, Leila and Sonia Aissa (2010). *Effective Capacity of Delay-Constrained Cognitive Radio in Nakagami Fading Channels*. IEEE Transactions on Wireless Communications, Vol. 9, No. 3, pp. 1054 – 1062

Musavian, Leila et al (2010). *Effective Capacity for Interference and Delay Constrained Cognitive Radio Relay Channels.* IEEE Transactions on Wireless Communications, Vol. 9, No. 3, pp. 1054 - 1062

Nakagami, Minoru (1958). *Statistical Methods in Radio Propagation*. Ed. W. C. Hoffman. Proceedings of a Symposium, University of California, June 18 - 20, 1958: Pergamon Press, NY, U.S.A

Pajala, Elina, et al., (2006). *An Improved Simulation Model for Nakagami-m Fading Channels for Satellite Positioning Applications*. Proceedings of the 3rd Workshop on Positioning, Navigation, and Communication, WPNC'06, pp. 81 - 90

Patzold, M. and Laue, F. (1998). *Statistical Properties of Jakes' Fading Channel Simulator.* Vehicular Technology Conference (VTC 98), 1998 IEEE 48th, Vol. 2, pp. 712 – 718

Pätzold, Matthias (2002). *Mobile Fading Channels*. John Wiley & Sons, Ltd, United Kingdom. ISBN 0471 49549 2

Peh, Edward and Ying-Chang Liang (2007). *Optimization of Cooperative Sensing in Cognitive Radio Networks.* Proceedings of IEEE WCNC, Hong Kong, Mar. 2007

Rappaport S., Theodore (1996). *Wireless Communications: Principles and Practice.* The Institute of Electrical and Electronics Engineers, Inc., Prentice Hall PRT, U.S.A.

Rappaport S., Theodore (2002). *Wireless Communications: Principles and Practice.* 2nd edition. Prentice Hall PRT, Prentice-Hall, Inc., Upper Saddle River, NJ, U.S.A.

Razavi, Behzad (2009). *Challenges in the Design of Cognitive Radios*. IEEE 2009 Custom

Integrated Circuits Conference (CICC).

Ross, M. Sheldon, (2010). *Introduction to Probability Models*. 10th edition. Elsevier Inc., U.S.A. ISBN 978-0-12-375686-2

Sklar, Bernard (2001). *Digital Communications: Fundamentals and Applications*. 2nd edition. Prentice Hall PTR, Upper Saddle River, NJ, U.S.A

Tang, Li, and Zhu Hongbo. (2003). *Analysis and Simulation of Nakagami Fading Channel with MATLAB*. Asia-Pacific Conference on Environmental Electromagnetic, CEEM 2003, Nov. 4 - 7, 2003, pp. 490 – 494.

Weisstein, Eric W. (2012) *Gamma Distribution*. MathWorld--A Wolfram Web Resource [Online].Available:http://mathworld.wolfram.com/GammaDistribution.html. Accessed:21/10/2012

Wu, Dapeng and Rohit Negi (2003). Effective Capacity: A Wireless Link Model for Support of Quality of Service. IEEE Transactions on Wireless Communications, Vol. 2, No. 4, pp 630 - 642

Yao, Yu-Dong and Asrar U. H. Sheikh (1993). *Evaluation of Channel Capacity in a Generalized Fading Channel*. IEEE Trans. Vehicular Technology, Vol. 39, pp. 134 – 137

Youssef et al., (1996). *Fade Statistics in Nakagami Fading Environments*. IEEE Xplore

APPENDIX I

PROOF OF EQUATION FOR THE ERGODIC CAPACITY
OF CR WITH MRC OVER NAKAGAMI FADING

This section is intended to show the derivation of the CR capacity per Hz given in equation 5.7. It is assumed that the two paths in the CR channel, that is, $SU_{tx} \rightarrow SU_{rx}$ and $SU_{tx} \rightarrow PU_{rx}$ experience Nakagami-m fading and the secondary receiver (SU_{rx}) has MRC implemented with K number of receiver antennas.

$$\frac{C_{erg}}{B} = E_{h,g}\left[\log\left(\frac{\gamma_o h}{g}\right)\right] =$$

$$= \int_0^\infty \int_{g/\gamma_o}^\infty \log\left(\frac{\gamma_o h}{g}\right) P_g(g) P_{hMRC}(h) \quad dh \, dg$$

$$= \int_0^\infty \int_{g/\gamma_o}^\infty \log\left(\frac{\gamma_o h}{g}\right) \frac{m_o{}^{m_o} g^{m_o-1}}{\Gamma(m_o)} e^{-m_o g} \frac{m^{Km} h^{Km-1}}{\Gamma(Km)} e^{-mh} \quad dh \, dg$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int_0^\infty g^{m_o-1} e^{-m_o g} \int_{g/\gamma_o}^\infty \log\left(\frac{\gamma_o h}{g}\right) h^{Km-1} e^{-mh} \quad dh \, dg$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int_0^\infty g^{m_o-1} e^{-m_o g} \int_1^\infty \log(t) * \left(\frac{g}{\gamma_o}\right)^{Km-1} t^{Km-1} e^{-m \cdot \frac{g}{\gamma_o} t} \frac{g}{\gamma_o} \quad dt \, dg$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int_0^\infty g^{m_o-1} e^{-m_o g} \left(g/\gamma_o\right)^{Km} J_{Km}\left(mg/\gamma_o\right) dg$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}} \int_0^\infty g^{m_o-1} e^{-m_o g} g^{Km} J_{Km}\left(mg/\gamma_o\right) dg$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}} \int_0^\infty g^{m_o-1+Km} e^{-m_o g} J_{Km}\left(mg/\gamma_o\right) dg$$

At this point, it is ideal to state that $\gamma_o = \frac{1}{\lambda N_o B}$ and $J_n(x)$ for n positive integers has been defined as thus:

$$J_n(x) = \frac{(n-1)!}{x^n} \sum_{k=0}^{n-1} \frac{\Gamma(k,x)}{k!}$$

The complementary incomplete gamma function $\Gamma(k, x)$ in the above expression can be defined as (Alouini & Andrea, 1999:1179):

$$\Gamma(k,x) = \int\limits_{x}^{\infty} t^{\alpha-1} e^{-t} \, dt$$

$$\frac{C_{erg}}{B} = \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}} \int\limits_{0}^{\infty} g^{m_o-1+Km} e^{-m_o g} \frac{(Km-1)!}{\left(\frac{mg}{\gamma_o}\right)^{Km}} \sum_{\alpha=0}^{Km-1} \frac{\Gamma\left(\alpha, \frac{mg}{\gamma_o}\right)}{\alpha!} dg$$

$$= \frac{\frac{m_o{}^{m_o} m^{Km}(Km-1)!}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}}}{\left(\frac{m}{\gamma_o}\right)^{Km}} \int\limits_{0}^{\infty} g^{m_o-1+Km} e^{-m_o g} \frac{1}{g^{Km}} \sum_{\alpha=0}^{Km-1} \frac{\Gamma\left(\alpha, \frac{mg}{\gamma_o}\right)}{\alpha!} dg$$

$$= \frac{\frac{m_o{}^{m_o} m^{Km}(Km-1)!}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}}}{\left(\frac{m}{\gamma_o}\right)^{Km}} \int\limits_{0}^{\infty} g^{m_o-1} e^{-m_o g}(Km-1)! \sum_{\alpha=0}^{Km-1} \frac{\Gamma\left(\alpha, \frac{mg}{\gamma_o}\right)}{\alpha!} dg$$

$$= \frac{\frac{m_o{}^{m_o} m^{Km}(Km-1)!}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}}}{\left(\frac{m}{\gamma_o}\right)^{Km}} \sum_{\alpha=0}^{Km-1} \frac{1}{\alpha!} \int\limits_{0}^{\infty} g^{m_o-1} e^{-m_o g} \Gamma\left(\alpha, \frac{mg}{\gamma_o}\right) dg$$

$$= \frac{\frac{m_o{}^{m_o} m^{Km}(Km-1)!}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}}}{\left(\frac{m}{\gamma_o}\right)^{Km}} \sum_{\alpha=0}^{Km-1} \frac{1}{\alpha!} \int\limits_{0}^{\infty} g^{m_o-1} e^{-m_o g} \Gamma\left(\alpha, \frac{mg}{\gamma_o}\right) dg =$$

Applying the integral approximation given in (*equation* 6.455 in Gradshteyn and Ryzhik, 2007: 657), the ergodic capacity of CR can be further expressed as follows:

$$\frac{C_{erg}}{B} = \frac{\frac{m_o{}^{m_o} m^{Km}(Km-1)!}{\Gamma(m_o)\Gamma(Km)\gamma_o{}^{Km}}}{\left(\frac{m}{\gamma_o}\right)^{Km}} \sum_{\alpha=0}^{Km-1} \frac{1}{\alpha!} \frac{\left(\frac{m}{\gamma_o}\right)^{\alpha} \Gamma(m_o+\alpha)}{m_o \left(\frac{m}{\gamma_o}+m_o\right)^{m_o+\alpha}} \; {}_2F_1\left(1, m_o+\alpha; m_o+1; \frac{m_o}{\frac{m}{\gamma_o}+m_o}\right)$$

APPENDIX II

PROOF OF POWER CONSTRAINT Q / (N$_o$B) IN EQUATION 5.8

$$\frac{Q}{N_o B} = E_{h,g}\left[\frac{P(h,g)g}{N_o B}\right] =$$

$$= \int\limits_{0}^{\infty} \int\limits_{0}^{\gamma_o h} \left(\gamma_o - \frac{g}{h}\right) P_g(g) P_{hMRC}(h) dg \, dh$$

$$= \int\limits_{0}^{\infty} \int\limits_{0}^{\gamma_o h} \left(\gamma_o - \frac{g}{h}\right) \frac{m_o{}^{m_o} g^{m_o-1}}{\Gamma(m_o)} e^{-m_o g} \; \frac{m^{Km} h^{Km-1}}{\Gamma(Km)} e^{-mh} dg \, dh$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int\limits_0^{\infty} \int\limits_0^{\gamma_o h} \left(\gamma_o - \frac{g}{h}\right) g^{m_o-1} e^{-m_o g} \ h^{Km-1} \ e^{-mh} \ dg \ dh$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int\limits_0^{\infty} h^{Km-1} e^{-mh} \int\limits_0^{\gamma_o h} \left(\gamma_o - \frac{g}{h}\right) g^{m_o-1} e^{-m_o g} \ dg \ dh$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int\limits_0^{\infty} h^{Km-1} e^{-mh} \int\limits_0^{\gamma_o h} \left(\gamma_o - \frac{g}{h}\right) g^{m_o-1} e^{-m_o g} \ dg \ dh$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int\limits_0^{\infty} h^{Km-1} e^{-mh} \left[\int\limits_0^{\gamma_o h} \gamma_o g^{m_o-1} e^{-m_o g} - \frac{1}{h} \int\limits_0^{\gamma_o h} g^{m_o-1} e^{-m_o g} \ dg \right] \ dh$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int\limits_0^{\infty} h^{Km-1} e^{-mh} \left[\int\limits_0^{\gamma_o h} \gamma_o g^{m_o-1} e^{-m_o g} - \frac{1}{h} \int\limits_0^{\gamma_o h} g^{m_o-1} e^{-m_o g} \ dg \right] \ dh$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \int\limits_0^{\infty} h^{Km-1} e^{-mh} \left[\gamma_o m_o{}^{-(m_o-1)-1} \gamma(m_o, m_o\gamma_o h) - \frac{1}{h} m_o{}^{-m_o-1} \gamma(m_o+1, m_o\gamma_o h) \right] dh =$$

$$= \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} m_o{}^{-m_o-1} \left[\frac{\gamma_o}{m_o} \int\limits_0^{\infty} h^{Km-1} e^{-mh} \gamma(m_o, m_o\gamma_o h) \ dh - \int\limits_0^{\infty} h^{Km-2} e^{-mh} \gamma(m_o+1, m_o\gamma_o h) \ dh \right]$$

With the aid of equation 6.455.2 in (Gradshteyn & Ryzhik, 2007: 657), the integral approximation can be obtained as:

$$\frac{Q}{N_0 B} = \frac{\frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)}}{m_o{}^{m_o+1}} \left[\frac{\gamma_o}{m_o} \frac{(m_o\gamma_o)^{m_o}\Gamma(Km+m_o)}{m_o(m_o\gamma_o+m)^{Km+m_o}} {}_2F_1\left(1, Km+m_o; m_o+1; \frac{m_o\gamma_o}{m_o\gamma_o+m}\right) \right.$$
$$\left. - \frac{(m_o\gamma_o)^{m_o+1}\Gamma(Km-1+m_o+1)}{(m_o+1)(m_o\gamma_o+m)^{Km+m_o}} {}_2F_1\left(1, Km+m_o; m_o+2; \frac{m_o\gamma_o}{m_o\gamma_o+m}\right) \right]$$

$$\frac{Q}{N_0 B} = \frac{\frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} m_o{}^{m_o+1} \gamma_o{}^{m_o+1}\Gamma(Km+m_o)}{m_o{}^{m_o+1}(m_o\gamma_o+m)^{Km+m_o}} \left[m_o{}^{-3} {}_2F_1\left(1, Km+m_o; m_o+1; \frac{m_o\gamma_o}{m_o\gamma_o+m}\right) \right.$$
$$\left. - \frac{1}{(m_o+1)} {}_2F_1\left(1, Km+m_o; m_o+2; \frac{m_o\gamma_o}{m_o\gamma_o+m}\right) \right]$$

$$\frac{Q}{N_0 B} = \frac{m_o{}^{m_o} m^{Km}}{\Gamma(m_o)\Gamma(Km)} \frac{\gamma_o{}^{m_o+1}\Gamma(Km+m_o)}{(m_o\gamma_o+m)^{Km+m_o}} \left[\frac{{}_2F_1\left(1, Km+m_o; m_o+1; \frac{m_o\gamma_o}{m_o\gamma_o+m}\right)}{m_o{}^3} \right.$$
$$\left. - \frac{1}{(m_o+1)} {}_2F_1\left(1, Km+m_o; m_o+2; \frac{m_o\gamma_o}{m_o\gamma_o+m}\right) \right]$$

Where $\gamma(\alpha, x)$ represents the incomplete gamma function defined as $\Gamma(k, x) = \int_x^\infty t^{\alpha-1} e^{-t} \, dt$ and $_2F_1(a, b; c; z)$ is the Gaussian Hypergeometric function (Duan et al., 2010:5; Alouini & Andrea, 1999:1179).

APPENDIX III

PROOF OF THE JOINT PROBABILITY OF THE CHANNEL POWER GAINS, *g* and *h*

This part shows the proof of the joint probability distribution $f_v(v)$ given in equation 5.18. Recall the PDFs of channel power gains *g and h* in equations 4.25 and 4.27 respectively. If *g* and *h* are Gamma variables with parameters *Km* and m_o, the ratio of variable *g* and *h* is a variable with a beta prime distribution and parameters *Km* and m_o (Weisstein, 2012). Using **Jacobian** method, the joint probability distribution was obtained as follows:

$$J\left(\frac{u, v}{x, y}\right) = \begin{vmatrix} 1 & 1 \\ 1/y & -x/y^2 \end{vmatrix} = -\frac{x+y}{y^2} = -\frac{(1+v)^2}{u}$$

$$u = vy + y, h = \frac{u}{1+v}, g = \frac{uv}{1+v}, v = \frac{x}{y} = \frac{g}{h}$$

$$f(u, v) = \frac{m^{Km} m_o{}^{m_o} g^{m_o-1} h^{Km-1}}{\Gamma(m_o)\Gamma(Km)} e^{-m_o g} e^{-mh} \left(\frac{u}{(1+v)^2}\right)$$

$$f(u, v) = \frac{m^{Km} m_o{}^{m_o}}{\Gamma(m_o)\Gamma(Km)} \left(\frac{uv}{1+v}\right)^{m_o} \left(\frac{u}{1+v}\right)^{Km-1} e^{-m_o\left(\frac{uv}{1+v}\right)} e^{-m\left(\frac{u}{1+v}\right)} \left(\frac{u}{(1+v)^2}\right)$$

$$= \frac{m^{Km} m_o{}^{m_o}}{\Gamma(m_o)\Gamma(Km)} u^{m_o-1+Km-1+1} v^{m_o-1} \left(\frac{1}{1+v}\right)^{m_o-1+Km-1+2} e^{-u\left(\frac{m_o v + m}{1+v}\right)}$$

$$f_v(v) = \int_0^\infty f_{u,v} \, du$$

$$f_v(v) = \frac{m^{Km} m_o{}^{m_o}}{\Gamma(m_o)\Gamma(Km)} \frac{v^{m_o-1}}{(1+v)^{m_o+Km}} \int_0^\infty u^{m_o+Km-1} e^{-u\left(\frac{m_o v + m}{1+v}\right)} \, du$$

With the help of equation 3.351.3 in (Gradshteyn & Ryzhik, 2007: 340), the integral approximation becomes:

$$f_v(v) = \frac{m^{Km} m_o{}^{m_o}}{\Gamma(m_o)\Gamma(Km)} \frac{v^{m_o-1}}{(1+v)^{m_o+Km}} (m_o + Km - 1)! \left(\frac{m_o v + m}{1+v}\right)^{-m_o-Km}$$

$$= \frac{m^{Km} m_o{}^{m_o}}{B(Km, \, m_o)} \frac{v^{m_o-1}}{(1+v)^{m_o+Km}} \frac{1}{(1+v)^{-m_o-Km}} \frac{1}{(m_o v + m)^{m_o+Km}} =$$

$$= \frac{Z^{Km}}{B(Km,\ m_o)} \frac{v^{m_o-1}}{(Z+v)^{m_o+Km}}$$

APPENDIX IV

PROOF OF POWER CONSTRAINT EQUATION IN 5.20:

$$\frac{Q}{NoB} = \int_0^\beta \left(\beta^{\frac{1}{1+\alpha}} v^{\frac{\alpha}{1+\alpha}} - v \right) f_v(v) =$$

$$\frac{Q}{NoB} = \int_0^\beta \beta^{\frac{1}{1+\alpha}} v^{\frac{\alpha}{1+\alpha}} f_v(v)\, dv - \int_0^\beta v\, f_v(v)\, dv =$$

$$I_1 = \int_0^\beta \beta^{\frac{1}{1+\alpha}} v^{\frac{\alpha}{1+\alpha}} f_v(v)\, dv$$

$$I_2 = \int_0^\beta v\, f_v(v)\, dv$$

$$I_1 = \int_0^\beta \beta^{\frac{1}{1+\alpha}} v^{\frac{\alpha}{1+\alpha}} f_v(v)\, dv = \int_0^\beta \beta^{\frac{1}{1+\alpha}} v^{\frac{\alpha}{1+\alpha}} \frac{Z^{Km}}{B(Km,\ m_o)} \frac{v^{m_o-1}}{(Z+v)^{m_o+Km}}\, dv$$

$$I_1 = \frac{\beta^{\frac{1}{1+\alpha}} Z^{Km}}{B(Km,\ m_o) Z^{m_o+Km}} \int_0^\beta \frac{v^{\frac{\alpha}{1+\alpha}+m_o-1}}{\left(1+\frac{1}{Z}v\right)^{m_o+Km}}\, dv$$

Using integral approximation of equation 3.194.1 in (Gradshteyn and Ryzhik, 2007: 315)

$$I_1 = \frac{\beta^{m_o}}{B(Km,\ m_o) Z^{m_o}\left(m_o+\frac{\alpha}{1+\alpha}\right)}\ {}_2F_1\left(Km+m_o;\ m_o+\frac{\alpha}{1+\alpha}\ ;\ m_o+\frac{\alpha}{1+\alpha}+1,-\frac{\beta}{Z}\right)$$

$$I_2 = -\int_0^\beta v \frac{Z^{Km}}{B(Km,\ m_o)} \frac{v^{m_o-1}}{(Z+v)^{m_o+Km}}\, dv$$

$$I_2 = -\frac{Z^{Km}}{B(Km,\ m_o) Z^{m_o+Km}} \int_0^\beta \frac{v^{m_o+1-1}}{\left(1+\frac{1}{Z}v\right)^{m_o+Km}}\, dv$$

$$I_2 = -\frac{Z^{Km}}{B(Km,\ m_o) Z^{m_o+Km}} \int_0^\beta \frac{v^{m_o+1-1}}{\left(1+\frac{1}{Z}v\right)^{m_o+Km}}\, dv$$

$$I_2 = -\frac{\beta^{m_o+1}}{B(Km, \, m_o)(m_o + 1)Z^{m_o}} \, {}_2F_1\left(Km \, + \, m_o; \, m_o + 1; \, m_o + 2, -\frac{\beta}{Z}\right)$$

Putting I_1 and I_2 together, equation 5.20 was obtained.

APPENDIX V

PROOF OF EQUATION (5.23) FOR THE EFFECTIVE CAPACITY OF CR WITH MRC OVER NAKAGAMI FADING

This section depicts the derivation of the effective capacity per Hz given in equation. It is assumed that the two paths in the CR channel $SU_{tx} \rightarrow SU_{rx}$ and $SU_{tx} \rightarrow PU_{rx}$ experience Nakagami-m fading and the secondary receiver (SU_{rx}) has MRC implemented with K number of receiver antennas.

$$\epsilon_v\left\{\left(1 + \left[\left(\frac{\beta}{v}\right)^{\frac{1}{1+\alpha}} - 1\right]^+\right)^{-\alpha}\right\} \qquad E_c = \int_0^\beta \frac{\beta^{-\frac{\alpha}{1+\alpha}}}{v^{-\frac{\alpha}{1+\alpha}}} f_v(v) \, dv + \int_\beta^\infty f_v(v) \, dv$$

$$E_{c1} = \int_0^\beta \frac{\beta^{-\frac{\alpha}{1+\alpha}}}{v^{-\frac{\alpha}{1+\alpha}}} \frac{Z^{Km}}{B(Km, \, m_o)} \frac{v^{m_o-1}}{(Z + v)^{m_o+Km}} \, dv$$

$$E_{c1} = \frac{Z^{Km}\beta^{-\frac{\alpha}{1+\alpha}}}{B(Km, \, m_o)Z^{m_o+Km}} \int_0^\beta \frac{v^{m_o+\frac{\alpha}{1+\alpha}-1}}{\left(1 + \frac{1}{Z}v\right)^{m_o+Km}} \, dv$$

Again, using integral approximation of equation 3.194.1 in (Gradshteyn and Ryzhik, 2007: 315)

$$E_{c1} = \frac{\beta^{m_o}}{B(Km, \, m_o)Z^{m_o}\left(m_o + \frac{\alpha}{1+\alpha}\right)} \, {}_2F_1\left(Km \, + \, m_o; \, m_o + \frac{\alpha}{1+\alpha} \, ; \, m_o + \frac{\alpha}{1+\alpha} + 1, -\frac{\beta}{Z}\right)$$

$$E_{c2} = \int_\beta^\infty f_v(v) \, dv = \int_\beta^\infty \frac{Z^{Km}}{B(Km, \, m_o)} \frac{v^{m_o-1}}{(Z + v)^{m_o+Km}} \, dv$$

$$= 1 - \left[\frac{Z^{Km}}{B(Km, \, m_o)Z^{m_o+Km}} \int_\beta^\infty \frac{v^{m_o-1}}{\left(1 + \frac{1}{Z}v\right)^{m_o+Km}} \, dv\right]$$

$$= 1 - \left[\frac{\beta^{m_o}}{B(Km, \, m_o)m_oZ^{m_o}} \, {}_2F_1\left(Km \, + \, m_o; \, m_o \, ; \, m_o + 1, -\frac{\beta}{Z}\right)\right]$$

Putting E_{c1} and E_{c2} together, the effective capacity in equation (5.23) was obtained.